

Empirical Issues *in* Syntax *and* Semantics 12

> Edited by

Christopher Pinon 2019 CSSP Paris http://www.cssp.cnrs.fr/eiss12/ ISSN 1769-7158

Contents

| Preface | iv |
|---|------------------|
| Anne Abeillé & Shrita Hassamal Sluicing in Mauritian: A Fragment-Based Analysis | 1–30 |
| Aixiu An & Anne Abeillé Number Agreement in French Binomials | 31–60 |
| Micky Daniels & Yael Greenberg Extreme Adjectives in Comparative Structures and <i>even</i> | 61–89 |
| Emilie Destruel & Joseph P. De Veaugh-Geiss (Non-)Exhaustivity in French <i>c'est</i> -Clefts | 91–120 |
| Regine Eckardt & Andrea Beltrama Evidentials and Questions | 121–155 |
| Yanwei Jin & Jean-Pierre Koenig Expletive Negation in English, French, and Mandarin: A Semantic and Language Production Model | 157–186 |
| Fabienne Martin & Zsófia Gyarmathy A Finer-grained Typology of Perfective Operators | 187–216 |
| Vítor Míguez On (Un)certainty: The Semantic Evolution of Galician <i>segurame</i> | 217–246 nte |
| Emanuel Quadros The Semantics of Possessive Noun Phrases and Temporal Modi | 247–276 fiers |



Preface

This is the twelfth volume of the series *Empirical Issues in Syntax and Semantics* (EISS), which, like the preceding eleven volumes of the series, is closely related to the conference series *Colloque de Syntaxe et Sémantique à Paris* (CSSP). The nine papers included in the present volume are based on presentations given at CSSP 2017, which took place on 23–25 November 2017 at École Normale Supérieure (http://www.cssp.cnrs.fr/cssp2017/index_en.html). CSSP 2017 had a small thematic session entitled *Discourse particles*, but since the number of papers from the thematic session submitted to the volume was low, they are not grouped separately.

I would like to take this opportunity to thank the reviewers, whose comments aided the authors in revising the first drafts of their papers, sometimes substantially. With their permission, the reviewers were (in alphabetical order by column):

| Berthold Crysmann | Caroline Heycock | Gerhard Schaden |
|---------------------|-------------------|------------------|
| Hans-Martin Gärtner | Alda Mari | Benjamin Spector |
| Jonathan Ginsburg | Laurent Roussarie | Jesse Tseng |

Finally, I would also like to thank both the scientific committee and the organizing committee of CSSP 2017 (http://www.cssp. cnrs.fr/cssp2017/contact/index_en.html) for their efforts in planning and organizing a memorable conference.

Christopher Pinon December 2019



Sluicing in Mauritian: A Fragment-Based Analysis

Anne Abeillé • Shrita Hassamal

Abstract Sluicing, since Ross 1969, has been a subject of syntactic debate: is the sluice a clause with full syntactic structure (Chung et al. 1995, Merchant 2001, among others) or just a reduced fragment (Ginzburg & Sag 2000, Culicover & Jackendoff 2005, among others)? We uncover some syntactic properties of sluices in Mauritian, a French-based creole, arguing for a fragment-based analysis.

Keywords sluicing \cdot ellipsis \cdot wh-word \cdot creole \cdot fragment \cdot pronoun

A. Abeillé, LLF/University of Paris 7-Diderot, abeille@linguist.univ-paris-diderot.fr S. Hassamal, Mauritius Institute of Education and LLF/University of Paris 7-Diderot, shritahassamal@gmail.com

In Christopher Pinon (Ed.), Empirical Issues in Syntax and Semantics 12, 1–30. Paris: CSSP. http://www.cssp.cnrs.fr/eiss12/ © 2019 Anne Abeillé & Shrita Hassamal

1 Introduction

1.1 What is Sluicing?

Sluicing is the name of a type of elliptical clause, with incomplete syntax and only a *wh*-phrase remnant, the full interpretation of which can be recovered from some antecedent in the context (Ross 1969). Sluicing can appear in the form of reduced interrogatives (1) or as short questions (2).

(1) a. John drinks something, but I don't know [what].

(Ross 1969)

- b. Someone solved the problem, but [who] is not clear.
- (2) a. Speaker 1: John's been drinking again? Speaker 2: What? (Culicover & Jackendoff 2005:47)
 - b. Speaker 1: I'm leaving soon. Speaker 2: When?

Sluicing can also occur as a reverse sluice, that is, in a construction where the *wh*-phrase precedes the full clause (3) or as an exophoric sluice, when the antecedent refers to the extralinguistic utterance context (4). Sprouting is another subtype of sluicing, where the *wh*-word has no explicit correlate in the matrix clause (or the preceding turn) (2a, 5).

- (3) I don't know [what], but he ate something bad.
- (4) [*Pointing in a shop*] "Could you tell me [how much]?"
- (5) John will go to Paris, but I don't know [when].

1.2 Two Syntactic Analyses of Sluices

The syntax of sluicing has been a matter of considerable debate and one distinguishes two main analyses:

- A deletion analysis (Ross 1969, Merchant 2001, among others)
- A fragment analysis (Ginzburg & Sag 2000, Culicover & Jackendoff 2005, among others)

Under the deletion analysis, an operation deletes the sentence (S or IP) that follows the *wh*-word (6). Since the sluice has a complete underlying structure, its semantics is directly computed from the syntax.

(6) John drinks something, but I don't know [what he's been drinking]_S.

Under the fragment analysis, the sluice is a base-generated fragment, which behaves externally as a clause (S or IP) but internally as a phrase (NP or DP) (7). The semantics is computed using information from the context.

(7) John drinks something, but I don't know $[[what]_{NP}]_S$.

Several arguments have been given in favour of the sentential status of sluices, which is common to both lines of analysis. As noted by Ross (1969), they occur after verbs that only take sentential complements, like *wonder* (8), and trigger singular agreement in subject position, unlike nominal subjects (9).

- (8) a. John is meeting someone tonight, I wonder [who].
 - b. I wonder [who John is going to meet tonight].
 - c. *I wonder the answer/the person ...
- (9) a. Some of these problems are solvable, but [which problems] is not obvious.
 - b. These problems are/*is not obvious.

As noted by Culicover & Jackendoff (2005:269), they trigger English particle placement like clausal complements (10b) and not like phrasal ones (10c).

- (10) a. He did something illegal but I never found out [what]/ *[what] out.
 - b. I never found out [what he did]/*[what he did] out.
 - c. He never found out [the answer]/[the answer] out.

The arguments in favour of a full-fledged syntactic structure of sluices, under a deletion analysis, on the other hand, seem to be weaker.

As noted by Ross (1969), the pro-forms are not necessarily the same in the antecedent and in the sluice. For example, one has to account for "sloppy" identity: the sluice means 'how to crane my neck' in (11a), and 'how to say I'm sorry' in (11b), as well as for 'vehicle change' (Sag & Nykiel 2011): the sluice means 'who is following you' in (11c).

- (11) a. Paul knows how to crane his neck but I don't know [how]. (Ross 1969)
 - b. I know how to say I'm sorry, and Bill knows [how], too.
 - c. Speaker 1: Someone is following me. Speaker 2: I wonder [who].

As noted by Hoeksema (2014), the Dutch *wh*-pronoun *waar* 'what' is also relevant here, as an R-pronoun which occurs to the left of preposition: it is obligatory in a *wh*-interrogative clause with preposition stranding (12a), but cannot appear in the corresponding sluice (12b), where the ordinary *wh*-pronoun *wat* 'what' appears.

- (12) a. Marie kijkt ergens naar, maar ik weet niet Marie looks somewhere at, but I know not waar/*wat zij naar kijkt. what.R/*what she at looks 'Marie looks at something but I don't know at what she looks.'
 b. Marie kijkt ergens naar maar ik weet niet
 - Marie kijkt ergens naar, maar ik weet niet Marie looks somewhere at, but I know not wat/*waar. what/*what.R

'Marie looks at something, but I don't know at what.'

As noted by Ross (1969) and Chung et al. (1995), island constraints, which apply to interrogative clauses, do not apply to sluices. (13) illustrates the Complex Noun Phrase Constraint, and (14), the Coordination Structure Constraint.

- (13) a. Bo talked to the people who discovered something, but we don't know [what].
 - b. *Bo talked to the people who discovered something, but we don't know [what Bo talked to the people who discovered].
- (14) a. Terry wrote an article about Lee and a book about someone else from East Texas, but we don't know [who].
 - b. *Terry wrote an article about Lee and a book about someone else from East Texas, but we don't know [who Terry wrote an article about Lee and a book about].

Merchant (2001) proposes that ellipsis may "repair" islands, but in a fragment analysis (7), the lack of island effects is actually predicted since there is no underlying syntactic structure for them to apply to.

Case assignment in sluices has been considered an argument in favour of an underlying verb, but Jacobson (2016) argues that, in fact, deletion-based approaches face intrinsic problems accounting for this. As pointed out by Ross (1969), in German, the sluice must be accusative if the antecedent verb is *sehen* 'see' (15a), and dative with an antecedent verb like *schmeicheln* 'flatter' (15b), which are known as "connectivity" effects. This is why Ginzburg & Sag (2000) and Culicover & Jackendoff (2005) in their fragment analysis proposed mechanisms for ensuring syntactic parallelism constraints between the fragment head (*wen* or *wem* 'who') and its salient correlate (*jemanden* or *jemandem* 'someone') (see section 3 below).

 (15) a. Er hat jemanden gesehen, ich weiss nicht he has someone-ACC seen, I know not wen/*wem. who-ACC/*DAT
 'He has seen someone, I don't know who.'

'He has seen someone, I don't know who.'

 b. Er will jemandem schmeicheln, ich weiss nicht, he wants someone-DAT flatter, I know not wem/*wen. who-DAT/*ACC
 (He wants to flatter someone I den't know who '

'He wants to flatter someone, I don't know who.'

However, as pointed out by Kim (2015), case-marking is not obligatory in Korean direct sluices, be it syntactic (16a) or semantic case (16b).

| (16) | a. | Speaker 1: | Mimi-ka | nwukwunka-lul | manna-ss-e. |
|------|----|-------------|-------------|--------------------|------------------|
| | | | Mimi-Noм | someone-ACC | meet-PST-DECL |
| | | | 'Mimi met | someone.' | |
| | | Speaker 2: | Nwukwu? | /Nwukwu-lul? | |
| | | | who/who- | ACC | |
| | b. | Speaker 1: | Mimi-ka | nwukwunka-lopy | wuthe senmwul-ul |
| | | | Mimi-Nom | someone-src | gift-ACC |
| | | pat-ass-e. | | | |
| | | receive-PST | -DECL | | |
| | | | 'Mimi recei | ved a gift from so | omeone.' |
| | | Speaker 2: | Nwukwu? | /Nwukwu-lopwu | the? |
| | | | who/who- | SRC | |
| | | | 'Who?/From | m whom?' | |
| | | | | | |

Preposition omission has been taken as another argument for a syntactic analysis of the sluice with an underlying full clause. According to Merchant's (2001) generalization, preposition omission under sluicing is possible only in those languages that allow for preposition stranding in regular *wh*-interrogatives (e.g., English, Norwegian). In this approach, there are two sources for (17a), one with preposition stranding (and preposition deletion) (17b), one with piedpiping (17c).

- (17) a. Peter was talking with someone, I don't know who/with whom.
 - b. I don't know who Peter was talking with.
 - c. I don't know with whom Peter was talking.

However, counterexamples exist in many languages, which allow preposition omission in sluices, but not preposition stranding: for example, Polish (Sag & Nykiel 2011), Romance (Fernández et al. 2007), Indonesian (Sato 2011), and Korean (Kim 2015). French allows for preposition omission in sluices (18a), but not in full interrogatives (18b). Notice that an alternative source, with clefting, as proposed by van Craenenbroeck (2010), is not possible in this case (18c).

- (18) a. Paul travaille sur quelque chose, je ne sais pas (sur) quoi.'Paul works on something, I don't know (on) what.'
 - b. Je ne sais pas sur quoi il travaille/*quoi il travaille sur.'I don't know with what he works/what he works with.'
 - c. *Paul travaille sur quelque chose, je ne sais pas quoi c'est/ c'est quoi.
 (Paul sur des en constituire de des (tobres en bet it is (it is))

'Paul works on something, I don't know what it is/it is what.'

Thus, a simple deletion-under-identity approach seems to face more problems than it solves. We show that Mauritian data favour a sentential analysis, while bringing more arguments against a deletion approach.

2 Mauritian Sluices

Mauritian is a French-based creole spoken in Mauritius, with no official status and a recent written form (Hookoomsing 2004, Carpooran 2011). We rely on Baker 1972, Henri 2010, Syea 2012, and Hassamal 2017 for its clausal syntax.

Little work has been done on ellipsis in creoles, with the exception of Costa et al. 2012, and we do not know of any on Mauritian. We rely on fieldwork surveys and on the few corpora available, such as the works of the Mauritian writer Dev Virahsawmy (https://boukiebanane.com/).

2.1 Verb-Form Alternation and Sluicing

The conjugation of Mauritian verbs displays two alternating forms: a long form and a short form (Baker 1972), with some 30% of verbs exhibiting syncretism. This morphological alternation (Bonami et al. 2011) depends on syntactic and discursive factors (Henri & Abeillé 2008, Henri 2010). For example, the short forms *manz* 'eat' and *koz* 'speak' are used before a canonical phrasal complement (19) and the long forms *manze* 'eat' and *koze* 'speak' are used otherwise (20) (Baker 1972, Henri 2010).¹

- (19) a. Mo'nn manz [pom] yer. 1SG-PERF eat.SF apple 'I ate apples yesterday.'
 - b. Mo koz [ek Pol].
 1SG speak.SF with Pol
 'I speak with Pol.'
- (20) a. Mo'nn manze/*manz yer. 1SG-PERF eat.LF/*SF yesterday 'I ate yesterday.'
 - b. Li'nn koze/*koz yer.
 35G-PERF speak.LF/*SF yesterday
 'He spoke yesterday.'

In *wh*-interrogatives, the *wh*-phrase is usually extracted in a leftperipheral position (21) and the long forms *manze* 'eat' and *koze* 'speak' are used. But it can also occur in situ, without an echo interpretation (Syea 2017) and the short forms *manz* 'eat' and *koz* 'speak' are used (22).

(i) Me Pol inn MANZE pom-la ! 'But Paul DID eat the apple!'

Some postverbal adverbs also trigger the short form and can be analysed as complements (Hassamal 2017, Hassamal et al. 2019):

(ii) Pol dans bien. Paul dances.sF well 'Paul dances well.'

¹The long form can also be used with a postverbal phrasal complement in the case of verum focus (Henri 2010). We ignore it here.

- (21) a. [Ki gato] to'nn manze/*manz yer ? which cake 2sg-PERF eat.LF/*sF yesterday 'Which cake did you eat yesterday?'
 - b. [Ek kisannla] to'nn koze/*koz ? with whom 2sG-PERF speak.LF/*sF
 'With whom did you speak?'
- (22) a. To'nn manz [ki gato]? 2SG-PERF eat.SF which cake Lit. 'You ate which cake?'
 - b. To'nn koz [ek kisannla]?2SG-PERF speak.SF with whom?Lit. 'You spoke with whom?'

Interestingly, the long form must occur before a sentential complement (23).

- (23) a. Mo panse/*pans [(ki) li pe ale].1SG think.LF/*SF (that) 3SG PROG go'I think that (s)he is leaving.'
 - b. Mo pa trouve/*trouv [si Pol pe koze].
 1SG NEG see.LF/*SF if Paul PROG speak
 'I do not see whether Paul is speaking.'

Like other subordinate clauses, a *wh*-interrogative clause triggers the long form of the main verb (24a), unless preceded by a phrasal complement (24b). On the other hand, if the phrasal complement follows the sentential complement, the matrix verb is in long form (24c). This is why Henri (2010) analyzes sentential complements as extraposed, and triggering the extraposition of following complements as well.

- (24) a. Pol inn manz gato, me mo pa'nn trouve/*trouv Paul PERF eat.SF cake, but 1SG NEG-PERF see.LF/*SF
 [ki gato li'nn manze]. which cake 3SG-PERF eat.LF
 'Paul has eaten a cake, but I did not see which cake he ate.'
 - b. Pol inn manz gato, mo demann [mwa] [ki gato Paul PERF eat.SF cake, 1SG ask.SF 1SG which cake li'nn manze].
 3SG-PERF eat.LF.

'Paul has eaten a cake, I wonder which cake he ate.'

c. Pol inn manz gato, mo demande/*demann [ki Paul PERF eat.SF cake, 1SG ask.LF/*SF which gato li'nn manze] [ar tou dimounn].
cake 3SG-PERF eat.LF to everybody.
'Paul has eaten a cake, and I ask everyone which cake he ate.'

Interestingly, a sluice also triggers the long form of the main verb (25a), unless preceded by another complement (25b).

(25) a. Pol inn manz gato, me mo pa'nn trouve/*trouv Paul PERF eat cake, but 1SG NEG-PERF see.LF/*SF
[ki gato]. which cake

'Paul has eaten a cake, but I did not see which cake.' Pol inn manz gato, me mo demann/*demande

b. Pol inn manz gato, me mo demann/*demande Paul PERF eat cake, but 1SG ask.LF/*SF
[mwa] [ki gato].
1SG which cake

'Paul has eaten a cake, but I wonder which cake.'

Hence, the sluice behaves like a clausal complement.

2.2 Wh-Words and Sluicing

As in French, most Mauritian *wh*-words can occur fronted (26) or *in situ* (27), for example, *kisannla* 'who', *kouma* 'how' and *kan* 'when', without an echo interpretation (Syea 2012).

- (26) a. Kisannla to'nn trouve ? who 2SG-PERF see.LF'Who did you see?'
 - b. Kan to pou manze?when 2SG FUT eat.LF'When will you eat?'
- (27) a. To'nn trouv kisannla ?2SG-PERF see.SF who?'Who did you see?'
 - b. To pou manze kan?2SG FUT eat.LF when'When will you eat?'

However, for the inanimate 'what', two forms, ki^2 and *kiete*, occur in a complementary distribution: ki must be fronted (28a), while *kiete* appears in situ (28b) or in isolation (28c).

(28) a. Ki/*Kiete to pou manze ? what 2SG FUT eat.LF 'What will you eat?'

²There are other uses for the form ki in Mauritian (Alleesaib 2008, Véronique 2007): the determiner ('which') (i), the complementizer ('that') (ii), and the discourse particle ('what?!!') (iii).

- (i) Ki gato to pou manze ? 'Which cake will you eat?'
- (ii) Li'nn dir ki li pou vini. 'He said that he will come.'
- (iii) Speaker 1: Mo'nn aret fime. 'I stopped smoking.' Speaker 2: Ki ?!! 'What?!!'

b. To pou manz kiete/*ki?
2SG FUT eat.SF what
'You will eat what?'
c. Speaker 1: Pol inn manz brinzel. Paul PERF eat.SF eggplant

'Paul ate eggplant.' Speaker 2: Kiete/#Ki ? 'What?'

Similarly, the locative 'where' has several forms in Mauritian: *kot*, *kote* and *kotsa*. *Kot* and *kote* are in complementary distribution: *kot* must be extracted (29a), while *kote* must be in situ (29b) or in isolation (29c).

- (29) a. Kot/*Kote to pou ale? where 2SG FUT gO.LF 'Where will you go?'
 - b. To pou al kote/*kot?2SG FUT go.SF where'You will go where?'
 - c. Speaker 1: Pol inn al deor. Paul PERF go.SF abroad 'Paul went abroad.' Speaker 2: Kote/*Kot ? 'Where?'

The third form *kotsa* is unconstrained and can occur as fronted (30a), in situ (30b) or in isolation (30c).

- (30) a. Kotsa to pou ale? where 2SG FUT gO.LF 'Where will you go?'
 - b. To pou al kotsa ?2SG FUT go.SF where 'You will go where?'

c. Speaker 1: Pol inn al deor. Paul PERF go.SF abroad 'Paul went abroad.' Speaker 2: Kotsa ? 'Where?'

Interestingly, only the in situ forms, *kiete* 'what' (31a) and *kote* 'where' (31b), are allowed in sluicing, whereas only the fronted forms *ki* (32a) and *kot* (32b) are allowed in full subordinate clauses.

 (31) a. Pol inn manz kitsoz, me mo pa'nn trouve Paul PERF eat.SF something, but 1SG NEG-PERF see.LF [kiete]/*[ki]. what
 'Paul ate something, but I didn't see what.'

 b. To'nn ferm li dan enn kaso, mo pa kone 2SG-PERF close.SF 3SG in IND jail, 1SG NEG know.LF [kote]/*[kot]. where

'You locked him/her in a jail, I do not know where.'

(Dev Virahsawmy, Toufann, Act 1, Scene 7)

- (32) a. Mo pa'nn trouve [ki/*kiete Pol inn manze]. 1SG NEG-PERF see.LF what Paul PERF eat.LF 'I didn't see what Paul ate.'
 - b. Mo pa kone [kot/*kote to'nn ferme li].
 1SG NEG know where 2SG-PERF lock 3SG
 'I do not know where you locked him/her.'

While use of the matrix verb long form with sluices is fully parallel to what we find with clausal complements, the data above challenge a deletion analysis with an underlying clause with a fronted *wh*-word (6). An underlying structure with a *wh*-word is not appealing, since, as English, Mauritian does not allow for an embedded interrogative clause with an in situ *wh*-element (33), outside echo uses.

- (33) a. *Mo pa'nn trouve, Pol inn manz kiete. 1SG NEG-PERF see.LF Paul PERF eat.SF what Intended: 'I didn't see what Paul has eaten.'
 - b. *Mo pa kone, Pol inn al kote. 1SG NEG know.LF, Paul PERF go.SF where Intended: 'I do not know where Paul went.'

Another tentative deletion analysis could be based on an underlying cleft construction, as proposed by van Craenenbroeck (2010) for Dutch. It is true that clefts only allow for strong *wh*-forms (34).

(34) a. Se kiete/*ki ki'nn kase? it what that-PERF break.LF 'It is what that has broken?'

b. Se kote/*kot ki to pe ale?it where that 2SG PROG gO.LF'It is where that you are going?'

But this line of analysis would not apply to indirect sluices:

- (35) a. *Mo pa'nn trouve, se kiete. 1SG NEG-PERF see.LF it what Intended: 'I didn't see what it is.'
 - b. *Mo pa kone, se kote.
 I NEG know it where
 Intended: 'I don't know where it is.'

It would also leave unexplained the possibility of preposition omission in sluices (36a), which is not possible in clefts (36c) nor in full *wh*-clauses (36b). Like French, Mauritian is a non-preposition stranding language (36d) (see also Sag & Nykiel 2011, Nykiel 2013 for a criticism of Merchant's (2001) generalization).

- (36) a. Pol inn koz ek enn kamarad, me mo pa Paul PERF speak.SF with IND friend, but ISG NEG kone (ek) kisannla. know.LF (with) who.
 'Paul has spoken with a friend but I do not know (with) whom.'
 b. Mo pa kone *(ek) kisannla li'nn koze.
 - b. Mo pa kone *(ek) kisannla li'nn koze. 1SG NEG know.LF *(with) whom 3SG-PERF speak.LF 'I do not know with whom (s)he spoke.'
 - c. Se *(ek) kisannla ki li'nn koze ?it *(with) whom that 3sg-PERF speak.LF'It is with whom that (s)he spoke?'
 - d. *Kisannla Pol inn koz/koze ek ?who Paul PERF speak.sF/LF with 'Who did Paul speak with?'

We thus conclude that the hybrid syntactic behaviour of Mauritian embedded sluices raises a challenge for syntactic analysis: on the one hand, they behave like clauses, triggering the matrix verb's long form; on the other hand, they only comprise in situ *wh*-words, unlike full embedded clauses. This may be a problem for deletion-based analyses.³ We show how it can be accounted for in a fragment-based approach.

- (i) Me ki li'nn manze ? 'But what did he eat?'
- (ii) Li'nn manz kiete exacteman ? 'What exactly did he eat?'

³As noted by a reviewer, a deletion analysis could use extra prosodic constraints. However, it is not the case that ki 'what' and kot 'where' are always sentence initial (i), nor are *kiete* 'what' and *kote* 'where' always sentence-final (ii), and the sluice is not necessarily preceded by a prosodic boundary. (We do not consider prosody any further here.)

3 An HPSG Analysis

To reconcile the apparently contradictory properties of Mauritian sluices, we use Ginzburg & Sag's (2000) approach, which includes fragment-like analyses for sluices and short answers: on this view, fragments are unary clauses with a full sentential meaning but a minimal syntactic structure reduced to a phrase. Thus, it is expected that sluices behave externally as clauses but internally as phrases:

- (37) a. Mo trouve [[kisannla]_{NP}]_S. 1SG see.LF who 'I see who (will come).'
 - b. Mo pa kone [[kote]_{Adv}]_S.
 1SG NEG know.LF where
 'I do not know where (he went).'
 - c. To'nn trouv [kisannla]_{NP}? 2SG-PERF see.SF who 'You saw who?'
 - d. Pol inn al [kot]_{Adv}? Paul PERF go.SF where 'Paul went where?'

Notice that *wh*-words used as sluices in (37a,b) differ from their ordinary use. When they are ordinary nominal or adverbial complements, they trigger the short form of the verb (37c,d).

3.1 Sluices as Fragments

Ginzburg & Sag (2000) rely on a cross-classification of clauses. Clauses come in different types (e.g., declarative or interrogative), and with different constituents (e.g., with a verbal head or not), as in figure 1.

Like verbal clauses, fragments may be declarative (for short answers) or interrogative (for short questions and sluices). Unlike full clauses, they are head-only clauses with a non-verbal head-daughter (DTRS) and a propositional content (CONT) inherited from the context (CTXT). Despite their non-verbal head, they have a verbal HEAD



Figure 1 Cross-classification of clauses

feature, like verbal clauses. We follow Ginzburg (2012) for the fragment's content: the propositional content is that of the maximal question under discussion (MAX-QUD), in which a variable x associated with the bare *wh*-phrase substitutes for the variable associated with its antecedent and is then abstracted away:



A sluice-clause inherits from interrogative clauses and from fragmentphrases (see figure 1):

(39) sluice-cl \Rightarrow interrog-cl & fragment-phr [DTRS $\langle [NON-LOC [WH \{ [INDEX [i]] \}]] \rangle]$

The head-daughter is coindexed with a salient constituent (SAL-UTT),

deor 'abroad' in (40), *kitsoz* 'something' in (41), with which it shares its syntactic features (CAT), hence the parallelism constraints.

- (40) Speaker 1: Pol inn al deor. Paul PERF go.SF abroad 'Paul went abroad.' Speaker 2: Kote/*Kot? 'Where?'
- (41) Pol inn manz kitsoz, me mo pa'nn trouve Paul PERF eat.SF something, but 1SG NEG-PERF see.LF kiete/*ki. what 'Paul ate something, but I didn't see what.'

As in English, the head-daughter of the fragment may be prepositional (42a), nominal (42b) or adverbial (42c). As the propositional content is inherited from the context (the previous clause or previous turn), they are interpreted as full clauses: 'With whom did Paul speak?' (42a), 'Who will come?' (42b), 'How many apples did Paul buy?' (42c).

(42) a. Speaker 1: Pol inn koz ek enn kamarad. Paul PERF speak.SF with IND friend 'Paul spoke with a friend.' Speaker 2: (Ek) kisannla? '(With) whom?'
b. Speaker 1: Bann dimounn pe vini. PLUR people PROG come.LF 'Some people are coming.' Speaker 2: Kisannla? 'Who?'
c. Speaker 1: Pol inn aste pom. Paul PERF buy.SF apple 'Paul bought apples.' Speaker 2: Komie? 'How many?'

In this analysis, preposition omission is handled as follows: the salient

constituent which serves as correlate can be the whole PP (*ek enn kamarad*) or the internal NP (*enn kamarad*), hence the two possible fragments in (42a).

The sharing of syntactic features ([2] in (38)) between the sluice and its correlate ensures syntactic parallelism. In case-marking languages, connectivity effects (section 1) are handled by the sharing of case features, which belong to CAT.

The sharing of syntactic features also prevents voice mismatches. In Mauritian, as in other languages, sluicing is out, as in (43b) where the sluice should be interpreted as passive with an active antecedent clause. Under our analysis, the fragment's daughter is a PP and its syntactic features fail to unify with those of the correlate *kikenn* 'someone', which is an NP in (43b).

- (43) a. Kikenn inn bat Zan, me mo pa kone [par Someone PERF hit John, but 1SG NEG know.LF by kisannla li'nn gagn bate]. whom 3SG-PERF get.SF hit.LF 'Someone hit John, but I don't know [by whom he was hit].'
 - b. *Kikenn inn bat Zan, me mo pa kone [par someone PERF hit.sF John, but 1sG NEG know.LF by kisannla].

Lit. 'Someone hit John, but I don't know [by whom].'

3.2 An Analysis of Verb-Form Alternation

As is standard in HPSG, we use a VFORM feature for verb forms (Sag et al. 2003). To account for verb alternation, we rely on two constraints on verb forms, leaving aside Verum Focus (Henri 2010): the short form (SF) requires a non-empty list (*nelist*) of complements (COMPS) while an empty list (*elist*) of complements triggers the long form (LF); VAL stands for VALENCE:

(44) Lexical constraints on verbs: $[\text{HEAD VFORM } SF] \Rightarrow [\text{VAL COMPS } nelist]$ $[\text{VAL COMPS } elist] \Rightarrow [\text{HEAD VFORM } LF]$

Following Henri 2010, a clausal complement is analysed as extraposed: it belongs to an EXTRA feature (Keller 1995, Crysmann 2003, Kay & Sag 2009) and does not appear on the verb COMPS list, hence the verb long form (46c). The same analysis applies to indirect sluices.

We use a SLASH feature to record which element has been extracted, and its value must unify with that of the filler in a longdistance dependency (Pollard & Sag 1994). Using Bouma et al.'s (2001) lexical analysis of extraction, words obey an argument conservation principle (45): an extracted complement is typed as noncanonical on the argument structure (ARG-ST) of the verb. It is not realized locally, and thus does not belong but to the list of complements: the verb with an extracted complement has thus an empty COMPS list, hence the long form (46b).

(45) Argument conservation principle:

$$word \Rightarrow \begin{bmatrix} valence \\ svbj 1 \\ spr 2 \\ comps 3 n-s-list \end{bmatrix}$$

$$arg-str 1 \oplus 2 \oplus 3 \oplus list (non-canon) \oplus 4$$

$$extra 4$$
with n-s-list = non-sentential list

A verb lexeme thus has different forms, depending on the realization of its syntactic arguments:

(46) a. koz 'speak' (no extraction)

$$\begin{bmatrix}
HEAD [VFORM SF] \\
VAL \begin{bmatrix}
SUBJ \langle 1 \rangle \\
COMPS \langle 2 \rangle
\end{bmatrix}$$
ARG-ST $\langle 1 NP[canon], 2 PP \begin{bmatrix} canon \\
FORM ek/avek \end{bmatrix} \rangle$

b. koze 'speak' (extracted complement)

$$\begin{bmatrix}
HEAD [VFORM LF] \\
VAL \begin{bmatrix}
SUBJ \langle 1 \rangle \\
COMPS \langle \rangle
\end{bmatrix}$$
ARG-ST $\langle 1 NP[canon], 2 PP \begin{bmatrix}
gap \\
FORM ek/avek
\end{bmatrix} \rangle$
SLASH {2}
c. kone 'know' (with a clausal complement)

$$\begin{bmatrix}
HEAD [VFORM LF] \\
VAL \begin{bmatrix}
SUBJ \langle 1 \rangle \\
COMPS \langle \rangle
\end{bmatrix}$$
ARG-ST $\langle 1 NP[canon], 2 S \rangle$
EXTRA {2}

3.3 Weak and Strong wh-Words

The observation that different *wh*-forms are used fronted and in situ may be a challenge for movement-based analyses of extraction (Chomsky 1977), but not necessarily for surface-oriented approaches (Sag & Fodor 1996).

We consider that the weak/strong distinction applies to Mauritian. Weak forms are not necessarily clitics (Cardinaletti & Starke 1999), but they cannot stand alone and have a more constrained distribution than strong forms. We consider *kot* 'where' and *ki* 'who' to be weak, since they cannot occur in isolation, and *kote* 'where' and *kiete* 'what' to be strong, while other *wh*-words (*kotsa* 'where', *kisannla* 'who', ...) are underspecified.

We consider that words can be weak, strong or underspecified, but that the distinction is not relevant for phrases. We thus define a three-valued WEAK feature, which is *boolean* (+ or -) for words and *non-applicable* (*na*) for phrases, as seen in figure 2.

We consider a general constraint that prevents [WEAK +] elements from being heads: this applies to fragments which are head-only phrases, and thus to sluices and fragment answers.



Figure 2 The values for WEAK

• *headed-phr* \Rightarrow Head-dtr [weak *na-or*-]

We consider additional constraints on Mauritian: weak forms can be fillers, while strong forms cannot.⁴ Notice that weak forms can also be subjects (47), and that with pied-piping, the strong forms are used, since they are complements of the preposition (48).

- (47) Ki/*Kiete pase la ? what happen.LF here 'What happened here?'
- (48) a. To'nn malad akoz kiete/*ki ? 2SG-PERF sick because what 'You fell sick because of what?
 - b. Akoz kiete/*ki to'nn malad ?
 because what 2sg-PERF sick
 'Because of what you fell sick?'

- (i) Pol inn trouv twa/*to. 'Paul has seen you.'
- (ii) Mo /*Mwa pou vini. 'I will come.'
- (iii) Speaker 1: Kisannla pou vini ? 'Who will come?' Speaker 2: Mwa/*Mo. 'Me.'

⁴Mauritian personal pronouns also display a weak/strong distinction (Syea & Véronique 2000): *mo* 'I' and *to* 'you' must be subjects (i), whereas *mwa* 'me' *twa* 'you' must be complements (ii) or occur in isolation (iii).

The weak forms ki 'what' and kot 'where' are [WEAK +] and the strong forms *kiete* 'what' and *kote* 'where' are [WEAK -] while other interrogatives are underspecified [WEAK *bool*].

In HPSG, interrogative words have a specific non-local feature (Pollard & Sag 1994), which marks interrogative clauses and ensures pied piping. Following Ginzburg & Sag (2000), we also use the non-local wH feature. Simplified lexical entries for *kot/kote* 'where' are in (49).



Subjects cannot be [WEAK -] (50a), while complements and extraposed constituents cannot be [WEAK +] (50b,c). Clefted elements, as other complements, must be strong.

- (50) a. hd-subj-phr \Rightarrow SUBJ-DTR [WEAK na-or+] b. hd-comps-phr \Rightarrow COMPS-DTR list ([WEAK na-or-])
 - c. hd-extra-phr \Rightarrow EXTRA-DTR list ([WEAK na-or-])

Wh-fillers, on the other hand, cannot be [WEAK -]. We consider it a constraint on interrogative clauses, while it may be a more general constraint on fillers:

(51) wh-inter-cl \Rightarrow FILLER-DTR [WEAK na-or+]

Since a full-fledged grammar of Mauritian is beyond the scope of



Figure 3 Kot to ale ? 'Where are you going?'

this paper, we provide only a few examples. Figure 3 shows an interrogative clause with a fronted *wh*-word: *Kot to ale*? 'Where are you going?'. This uses *hd-filler-phr* (figure 1), and the filler *kot* 'where' is [WEAK +]. It also uses *hd-subj-phr* and the subject *to* 'you' is [WEAK +], and the verb is not followed by a canonical complement (its COMPS list is empty), hence the long form of the verb *ale* 'go'. Following Bouma et al. (2001), there is no empty category and the extracted complement is noted in the SLASH value of the verb, and it is percolated up the syntactic tree until it is unified with the local features of the filler.

Figure 4 shows an interrogative clause with an in situ *wh*-word: *To al kote*? 'You go where?'. As in figure 3, it uses *hd-subj-phr* and the subject word *to* 'you' is [WEAK +]. It also uses *hd-comps-phr* and the complement (*kote* 'where') is [WEAK -]. The verb is followed by a complement (its COMPS list is not empty), hence the short form of the verb *al* 'go'.

Figure 5 shows a declarative clause with a sluice complement: *Zan kone kote* 'John knows where'. As a sentential complement, the sluice is extraposed. As a head-only phrase, its head cannot be weak, and the *wh*-word *kote* 'where' is [WEAK -]. The matrix verb is not followed by a canonical complement, and its form is long: *kone* 'know'.

4 Conclusion

We have shown that Mauritian sluices exhibit hybrid syntactic properties. On the one hand, they do not trigger the short form of the matrix verb, like clausal complements and unlike phrasal complements. On the other hand, they do not have the internal structure of an interrogative clause and only comprise in situ *wh*-words. These properties challenge a syntactic deletion analysis of sluices. We show that they can be handled by a HPSG grammar following Ginzburg & Sag's (2000) fragment-like analysis.

Acknowledgments This work was supported by strand 4 of the Laboratoire d'excellence Empirical Foundations of Linguistics (ANR-10-LABX-0083). We thank B. Crysmann, J. Ginzburg, F. Henri, the audiences of the GdRi



Figure 4 To al kote ? 'You're going where?'



Figure 5 Zan kone kote 'John knows where'

SEEPiCLa in Orléans (October 2017), the GRGC in March 2018 (Paris 8), the CSSP 2017 conference in Paris and the CIEC in Mahé (November 2018), as well as an EISS reviewer for their helpful comments. Finally, we thank L. Roussarie and the editor for their editorial assistance.

References

Alleesaib, M. (2008). *Se-ki* en créole mauricien : l'émergence d'un pronom. *Modèles linguistiques*, *57*, 11–33.

URL http://journals.openedition.org/ml/258

Baker, P. (1972). Kreol: A Description of Mauritian Creole. London: C. Hurst.

- Bonami, O., Henri, F., & Luís, A. R. (2011). The emergence of morphomic structure in Romance-based creoles. In *20th International Conference on Historical Linguistics*. Osaka.
- Bouma, G., Malouf, R., & Sag, I. A. (2001). Satisfying constraints on extraction and adjunction. *Natural Language and Linguistic Theory*, 19, 41–79.
- Cardinaletti, A., & Starke, M. (1999). The typology of structural deficiency: A case study of the three classes of pronouns. In H. van Riemsdijk (Ed.) *Clitics in the Languages of Europe*, (pp. 145–234). Berlin: Mouton de Gruyter.
- Carpooran, A. (2011). *Lortograf Kreol Morisien*. Phoenix, Mauritius: Akademi Kreol Morisien.

URL https://tinyurl.com/uqlglo4

- Chomsky, N. (1977). On *wh*-movement. In P. W. Culicover, T. Wasow, & A. Akmajian (Eds.) *Formal Syntax*, (pp. 71–132). New York: Academic Press.
- Chung, S., Ladusaw, W. A., & McCloskey, J. (1995). Sluicing and logical form. *Natural Language Semantics*, *3*, 239–282.
- Costa, J., Martins, A. M., & Pratas, F. (2012). VP ellipsis: New evidence from Capeverdean. In I. Franco, S. Lusini, & A. Saab (Eds.) Romance Languages and Linguistic Theory 2010: Selected Papers from 'Going Romance' Leiden 2010, (pp. 155–176). John Benjamins.
- Crysmann, B. (2003). On the efficient implementation of German verb placement in HPSG. In *Proceedings of RANLP 2003*, (pp. 112–116). Borovets, Bulgaria.
- Culicover, P. W., & Jackendoff, R. (2005). *Simpler Syntax*. Oxford University Press.

- Fernández, R., Ginzburg, J., & Lappin, S. (2007). Classifying non-sentential utterances in dialogue: A machine learning approach. *Computational Linguistics*, 33, 397–427.
- Ginzburg, J. (2012). The Interactive Stance. Oxford University Press.
- Ginzburg, J., & Sag, I. A. (2000). *Interrogative Investigations: The Form, Meaning, and Use of English Interrogatives*. Stanford, CA: CSLI Publications.
- Hassamal, S. (2017). *Grammar of Mauritian Adverbs*. Ph.D. thesis, University of Paris 7-Diderot.
- Hassamal, S., Abeillé, A., & Henri, F. (2019). Les adverbes en mauricien. *Faits de Langues*, 49(1), 69–98.
- Henri, F. (2010). A Constraint-Based Approach to Verbal Constructions in Mauritian: Morphological, Syntactic and Discourse-Based Aspects. Ph.D. thesis, University of Paris 7-Diderot.
- Henri, F., & Abeillé, A. (2008). Verb form alternations in Mauritian. In S. Müller (Ed.) Proceedings of the 15th International Conference on Head-Driven Phrase Structure Grammar, National Institute of Information and Communications Technology, Keihanna, (pp. 378–398). Stanford, CA: CSLI Publications.

URL https://tinyurl.com/utfgpp2

Hoeksema, J. (2014). Sluicing in Dutch: A problem for PF-deletion approaches. *SKASE Journal of Theoretical Linguistics*, 11(2).

URL http://www.skase.sk/Volumes/JTL26/pdf_doc/02.pdf

- Hookoomsing, V. (2004). *A Harmonised Writing System for the Mauritian Creole Language: Grafi Larmoni*. Port Louis, Mauritius: Ministry of Education and Scientific Research.
- Jacobson, P. (2016). The short answer: Implications for direct compositionality (and vice versa). *Language*, *92*, 331–375.
- Kay, P., & Sag, I. A. (2009). How hard a problem would this be to solve? In S. Müller (Ed.) Proceedings of the 16th International Conference on Head-Driven Phrase Structure Grammar, Georg-August-Universität Göttingen, Germany, (pp. 171–191). Stanford, CA: CSLI Publications. URL https://tinyurl.com/remqm2t
- Keller, F. (1995). Towards an account of extraposition in HPSG. In Seventh Conference of the European Chapter of the Association for Computational Linguistics, (pp. 301–306). Dublin, Ireland: Association for Computational Linguistics.

URL https://www.aclweb.org/anthology/E95-1045/

- Kim, J.-B. (2015). Syntactic and semantic identity in korean sluicing: A direct interpretation approach. *Lingua*, *166*, 260–293.
- Merchant, J. (2001). *The Syntax of Silence: Sluicing, Islands, and the Theory of Ellipsis*. Oxford University Press.
- Nykiel, J. (2013). Clefts and preposition omission under sluicing. *Lingua*, 123, 74–117.
- Pollard, C., & Sag, I. A. (1994). *Head-Driven Phrase Structure Grammar*. The University of Chicago Press.
- Ross, J. (1969). Guess who? In *Chicago Linguistic Society*, vol. 5, (pp. 252–286).
- Sag, I. A., & Fodor, J. D. (1996). Une analyse sans catégorie vide des phénomènes d'extraction. *Langages*, 122, 8–31.
- Sag, I. A., & Nykiel, J. (2011). Remarks on sluicing. In S. Müller (Ed.) Proceedings of the 18th International Conference on Head-Driven Phrase Structure Grammar, University of Washington, (pp. 188–208). Stanford, CA: CSLI Publications.

URL https://tinyurl.com/qohle9v

- Sag, I. A., Wasow, T., & Bender, E. M. (2003). *Syntactic theory: A Formal Introduction*. Stanford, CA: CSLI Publications, 2nd ed.
- Sato, Y. (2011). P-stranding under sluicing and repair by ellipsis: Why is Indonesian (not) special? *Journal of East Asian Linguistics*, *20*, 339–382.
- Syea, A. (2012). *The Syntax of Mauritian Creole*. London: Bloomsbury.
- Syea, A. (2017). *French Creoles: A Comprehensive and Comparative Grammar*. Routledge.
- Syea, A., & Véronique, G. D. (2000). L'absence de pronoms clitiques en créole mauricien. Langages, 34(138), 70–88. URL https://tinyurl.com/uybawqx

van Craenenbroeck, J. (2010). The Syntax of Ellipsis: Evidence from Dutch

- Dialects. Oxford University Press.
- Véronique, D. (2007). Ki en haitien, en mauricien et dans quelques autres créoles français. In M. Charolles, N. Fournier, C. Fuchs, & F. Lefeuvre (Eds.) Parcours de la phrase : Mélanges offerts à Pierre Le Goffic. Gap: Éditions Ophrys.

Number Agreement in French Binomials

Aixiu An • Anne Abeillé

Abstract Number agreement in determiner plus bare-noun coordination (D N & N) varies across languages. This paper presents an empirical study of French in which the construction has been claimed to only be felicitous with plural nouns (e.g., Heycock & Zamparelli 2005, Le Bruyn & de Swart 2014). We use large-corpus data and ran two acceptability rating experiments, showing that singular split binomials allow for two agreement strategies for the shared determiner: Closest Conjunct Agreement (*votre nom et prénom* 'your.sc name and first name' or Resolution (*vos nom et prénom* 'your.PL name and first name'). We also show that speakers' preferences are sensitive to noun animacy and the syntactic function of the binomial. We assume that animate binomials give rise to a potential ambiguity between a joint and a split reading. We propose an HPSG analysis that assigns the same syntactic structure to split and joint binomials, while allowing for different agreement strategies.

Keywords number agreement · binominals · split reading · animacy · French · large corpora · experimental syntax

```
A. An, LLF/University of Paris 7-Diderot, aixiu0806@gmail.com
A. Abeillé, LLF/University of Paris 7-Diderot, abeille@linguist.univ-paris-diderot.fr
```

```
In Christopher Pinon (Ed.), Empirical Issues in Syntax and Semantics 12, 31–60. Paris: CSSP. 
http://www.cssp.cnrs.fr/eiss12/
© 2019 Aixiu An & Anne Abeillé
```

1 Introduction

In French, determiners and nouns agree in number.¹ Singular nouns require singular determiners and plural nouns require plural determiners (1a). And two coordinated singular NPs (with *et* 'and') trigger plural agreement (1b).

¹French determiners also agree in gender. We ignore gender agreement here, but see An & Abeillé 2017 and Abeillé et al. 2018.

- (1) a. un chien /des chiens a.sG dog.sG /a.PL dog.PL
 - b. [un chien et un chat] intelligents a.sG dog.sG and a.sG cat.sG clever.PL

Crosslinguistically, binomials display specific agreement properties. When two bare nouns are conjoined with a shared determiner (D N1 & N2), they can be coreferent (joint interpretation) (2a) or not (split interpretation) (2b) (e.g., King & Dalrymple 2004, Heycock & Zamparelli 2005, Le Bruyn & de Swart 2014).

- (2) a. A friend and colleague is coming today.
 - b. My father and mother are separated.

In most languages, the determiner (D) agrees in number with both conjuncts in the joint interpretation, since they share the same index. However, languages have different strategies for the split interpretation: in English, as in Finnish and Hindi, the D is singular with singular nouns (3a), and nouns with different numbers cannot be coordinated with a shared D (3b), (3c) (King & Dalrymple 2004).

- (3) a. This boy and girl are eating a pizza.
 - b. *these boys and girl
 - c. *this boy and girls

Other languages also use a singular D but allow for conjuncts with different numbers. In that case, Spanish (4a) and Portuguese (4b), for example, allow Closest Conjunct Agreement (CCA). The shared D agrees with the first noun.²

(4) a. Los ataques de la aviación y the attacks of the.FSG aeroplanes.FSG and helicópteros rusos fueron constantes. helicopters.MSG Russian.MPL be.PST.3PL constant.PL

²In the glosses, FSG represents singular feminin gender, MPL represents plural maculin gender, and NPL represents plural neutral gender.
'The attacks of the Russian aeroplanes and helicopters were constant.' (Demonte & Perez-Jimenez 2012)
b. O presidente e amigo/amigos comeram the.MsG president.MsG and friend.MsG/PL ate.PL juntos. together 'The president and friend/friends ate together.' (Villavicencio et al. 2005)

CCA is also attested in Serbo-Croatian (Wechsler & Zlatić 2003, Dalrymple & Hristov 2010). In (5), the prenominal Adj agrees with the first conjunct and the participle agrees with the second conjunct.

(5) Nova kaŭca i kola su kŏstal-a puno. new.FSG house.FSG and car.NPL AUX.PL cost.NPL a.lot 'The new house and car cost a lot.' (Dalrymple & Hristov 2010)

Le Bruyn & de Swart (2014) assume that English-like and Spanishlike languages work similarly, since a plural D is not possible (4a), (6a), except for some rare cases with proper names ((6b) from Demonte & Perez-Jimenez 2012). They note that English may allow a mismatch of number in an appropriate context, as in (6c).

- (6) a. *These boy and girl are eating a pizza.
 - b. ... los temibles Capirucho y Capirote ...
 - ... the.MPL fearsome.PL Capirucho and Capirote ...
 - '... the fearsome Capirucho and Capirote '
 - c. These children and mother were living on charity of good people.

Assuming a hierarchical structure for coordination (Kayne 1994, Borsley 2005), examples (4a) and (5) can also be considered as cases of highest conjunct agreement (e.g., Marušič et al. 2007, Murphy & Puškar 2018): agreement with the first noun, which is hierarchically highest, or closer to D in terms of number of intervening nodes. According to Le Bruyn & de Swart (2014), Russian is one of the "typologically rare languages" that allows for a plural D with a coordination of singular nouns (King & Dalrymple 2004) (7a), and French is supposed to forbid both a singular D and a plural D (7b), except in some "frozen" cases such as *vos père et mère* ('your.PL father and mother').

(7) a. èti mužčina i ženšžina this.PL man.MSG and woman.FSG 'these man and woman' (King & Dalrymple 2004:95)
b. *Ce/*Ces marin et soldat sont souvent ensemble. this.SG/PL sailor and soldier are.PL often together (Heycock & Zamparelli 2005:3)

Contrary to Heycock & Zamparelli (2005) and Le Bruyn & de Swart (2014), we have found that both a singular and a plural D are permitted in French with two non-coreferent singular nouns (An & Abeillé 2017), as in (8).

(8) Indiquez votre/vos nom et prénom. indicate your.sg/.PL last name.MsG and first name.MsG 'Indicate your first and last name.'

However, this is not felicitous with all binomials. For instance, a plural D is better in (9a), and a singular D in (9b), respectively.

(9)frère d'Emmanuel Les/?Le a. et sœur the.PL/?.MSG brother.MSG and sister.FSG of Emmanuel Macron n'ont pas souhaité être exposés Macron NEG-AUX NEG want.PST-PT be exposed.MPL durant cette campagne. during this campaign 'The brother and sister of Emmanuel Macron didn't want to be exposed during this campaign.' (FrWAC, www.gala. fr)

 b. La/?Les simplicité et beauté sont pour moi the.FsG/.PL simplicity.FsG and beauty.FsG are for me étroitement associées. closely associated.PST-PT.FPL 'The simplicity and beauty are closely associated for me.' (FrWAC, www.auto-evasion.com)

We present an empirical study based on large-scale corpus data and two controlled experiments, investigating D number agreement with French singular binomials, as well as its interactions with other factors, such as noun animacy and syntactic function. The article is organized as followed: in section 2, we discuss the backgrounds of our study, showing that the ambiguity between joint and split readings may come from frequency and context effects. With corpus data in section 3, we show that non-human nouns are more frequent with a split reading while human nouns favor a joint interpretation. With split singular binomials, a plural D is preferred for human nouns and a singular D for non-human nouns. In section 4, we report two experimental studies, testing the acceptability of singular/plural D with human and non-human nouns, in two different positions: subject and object. In section 5, we compare our corpus data and our experimental results and propose a formal analysis in HPSG.

2 Research Questions

2.1 Factors favoring Closest Conjunct Agreement

According to Corbett (1991), the factors favoring Closest Conjunct agreement (singular) or agreement resolution (plural) crosslinguistically, involve the agreement controller (the element which governs agreement) and the agreement target (the element which marks agreement). In a corpus study of Medieval Spanish, German, Russian and Serbo-Croatian, in which the predicate can be either singular or plural with two singular conjoined nouns, Corbett counted agreement examples in two possible orders, subject-predicate and predicate-subject, and showed that resolution (plural) was found more frequently with animate nouns and in subject-predicate order.

Corbett also claims that the choice between resolution and nonresolution obeys an agreement hierarchy (10) (Corbett 1991). As we move rightwards along the hierarchy of targets, the likelihood of singular agreement decreases and that of resolution (plural) increases.

(10) attributive > predicative > relative pronoun > personal pronoun \leftarrow non-resolution resolution \rightarrow

He does not mention determiner agreement. Since Determiner-Noun order is fixed, we predict that the only factor that interacts with agreement is animacy. Amongst the interactions between number agreement and noun animacy, we predict that a plural D will be favored by animate nouns over non animates.

2.2 Binomials and Semantic Ambiguity

Binominal constructions are ambiguous between a joint (coreferent) and a split (non coreferent) reading. We observe that human nouns are often role and function names, and thus more compatible with a joint reading than non-human nouns, referring to objects, properties or situations. The plural animate binominal is ambiguous between a joint reading and a split reading in (11b), while the inanimate binominal does not involve this ambiguity (11b).

- (11) a. J' ai vu des collègues et amis I see.PST a.PL colleagues and friends 'I saw some colleagues and friends.' (ambiguous)
 b. J' ai vu des cahiers et crayons
 - I see.PST a.PL notebook.MPL and pencil.MPL 'I saw some notebooks ans pencils.' (split)

With singular nouns, both readings are possible with a singular D, while a plural D forces a split reading. To avoid this ambiguity, we suppose that singular animate Nouns favor a division of labor between plural D (split reading) (12a) and singular D (joint reading) (12b).

- (12) a. des collègue et ami 'a.PL colleague and friend.'
 - b. un collègue et ami'a.sG colleague and friend.'

We will examine whether a biased context favoring the split reading will affect the acceptability of a plural D and section 4.

It should be noted that animacy is rather a continuum than a binary distinction between animate-inanimate (Comrie 1989, Croft 2002, Zaenen et al. 2004, Haspelmath 2013). In this study, we distinguish humans and non-humans for the reason that coreference is more likely to occur with human nouns. Animals are considered as non-humans for the same reason (*des chiens et chats* ('dogs and cats') cannot have a joint reading).

We begin with corpus data comparing the distribution of human and non-human binomials, as well as D agreement. Then, we report two experimental studies controlling syntactic and semantic factors, in order to analyze the different factors affecting the acceptability of French binomials.

3 Corpus study

3.1 Methodology

We use the web-based corpus FrWac (1.6 billion words, Baroni et al. 2009). We found 371 000 tokens with the request $D \ N \ et \ N$. We annotated the number of nouns and determiners with Flemm (Namer 2000). Table 1 reports the numbers of different types of binomials.

The plural binomials are the most frequent ones. If N1 and N2 are both singular, D can be either singular ((13), (14)) or plural (15). With Dsg, the binominal can have either a joint (13) or a split reading (14).

(13) Le chanteur et poète québécois Gilles Vigneault the.msg singer.msg and poet.sg Quebec.m Gilles Vigneault

| construction | types | tokens | |
|----------------|--------|--------|--|
| Dsg Nsg et Nsg | 31412 | 51711 | |
| Dpl Nsg et Nsg | 1308 | 5137 | |
| Dsg Nsg et Npl | 5742 | 9490 | |
| Dpl Nsg et Npl | 724 | 1432 | |
| Dpl Npl et Nsg | 7586 | 13460 | |
| Dpl Npl et Npl | 55269 | 201503 | |
| total | 102041 | 282733 | |



publie en France un livre d' entretiens. publish.3sG in France a book of interviews 'The Quebec singer and poet, Gilles Vigneault, publishes a book of interviews in France.' (FrWAC, republique-des-lettres.fr)

- (14) Présentez -vous à la date et lieu introduce.imp yourself at the.FSG date.FSG and place.MSG indiqué pour suivre votre formation. indicated.MSG to follow.inf your training. 'Introduce yourself at the date and place indicated to follow your training.' (FrWAC, secours57.fr)
- (15) Les lieu et programme seront précisés the.PL place.MSG and program.MSG be.FUT.3PL specified.MPL sur le bulletin. on the bulletin 'The place and program will be specified on the bulletin.' (FrWAC, rao.free.fr)

We also found cases with a number mismatch. When N1 is singular and N2 is plural, Dpl (16a) is less frequent than Dsg (16b), which can be considered an instance of Closest Conjunct Agreement, or highest conjunct agreement.

- (16) a. L' atelier est fermé le dimanche et the workshop AUX. close.PST the.MSG Sunday and jours fériés. holiday.PL 'The workshop is closed on Sundays and public holidays.' (FrWAC, fram.fr)
 b. Cotta publication comportation pomototical
 - b. Cette publication comporte les nom et This.FSG publication contains the.PL last name and prénoms du débiteur. first name of.MSG debtor 'This publication contains the first and last names of the debtor.' (FrWAC, courdecassation.fr)

When N1 is plural and N2 singular, D is plural(17), which can be considered as a case of Closest Conjunct Agreement or resolution.³

(17) Voici pour les frères et sœur des photos here is for the.PL brother.FSG and sister.FSG some picture faites tout exprès pour eux. do.PST.FPL all especially for them 'Here are pictures made especially for the brothers and sisters.' (FrWAC, catherine-de-mercueil.over-blog.fr)

3.2 Animacy and Semantic Readings

We extracted the binomials with more than five occurrences (22600 tokens) and removed the errors (10640 tokens left).⁴ We annotated noun animacy with an external dictionary (Olivier Bonami, pers. comm.) and the joint or split reading manually. Table 2 reports the

³In FrWac, 88 cases were labeled as Dsg Npl and Npl, but all involve obvious errors. Thus, they are not listed in table 1.

⁴There are three major sources of errors: first, quite a lot of repeated tokens are found; second, some constructions which are in the form of *D N*¹ *et N*² are not binomials, like *une heure et quart* (one.F hour.FSG and half.MSG); another major kind of errors is that the two bare nouns are not in the same constituent, like *mise sur le marché et prise en charge des vaccins* (placing.FSG on the.MSG market.MSG and taking.FSG care of the vaccines).

| | joint | | split | | total | |
|------------|-------|--------|-------|--------|-------|--------|
| | types | tokens | types | tokens | types | tokens |
| humans | 196 | 2304 | 6 | 105 | 202 | 2409 |
| non-humans | 3 | 31 | 492 | 10535 | 495 | 10566 |
| total | 199 | 2335 | 498 | 10640 | 697 | 12975 |

 Table 2 Readings of singular binomials in FrWAC

number of tokens with the joint/split reading for human and nonhuman nouns. For human nouns, 97.03% of the examples have a joint reading(13), while for non-human nouns, only 0.6% (18).

(18) Le restaurant et bar Starlight propose un the.MSG restaurant.MSG and bar.MSG Starlight offer.3SG a menu international. menu international 'The restaurant and bar Starlight offers an international menu.' (FrWAC, expedia.fr)

The results illustrate important frequency differences between human and non-human nouns regarding the interpretation. Human nouns favor the joint reading while non-human nouns favour the split reading. We can infer that human nouns may have more comprehension difficulties for the split reading because of the spurious ambiguity, which may lower its acceptability.

3.3 Animacy and Number Agreement

Then, we now examine the agreement strategies for the split reading with respect to human and non-human nouns. Our results (table 2) report more examples with a plural D for human nouns (19), even if both singular and plural D are quite rare. However, for non-human nouns, the singular D is more frequent (see (14) and (15)).

| | Dsg | | Dpl | | total | |
|-----------|-------|--------|-------|--------|-------|--------|
| | types | tokens | types | tokens | types | tokens |
| human | 1 | 6 | 5 | 99 | 6 | 105 |
| non-human | 439 | 7507 | 53 | 2997 | 492 | 10535 |
| total | 440 | 7513 | 58 | 3096 | 498 | 10640 |

 Table 3 Dsg/Dpl with singular binomials in FrWAC

(19)

a. Les mari et femme sont d'accord the.PL husband.MSG and wife.FSG be.3PL of agreement sur le partage des biens. on the division of.PL property.PL 'The husband and wife agree on the division of these properties.' (FrWAC, judiciaire.blog.20minutes. fr)
b. Elle part loin à la recherche de nourriture pour che leguns for to the recent of food for

 b. Elle part loin à la recherche de nourriture pour she leaves far to the research of food for son mari et bébé.
 her.MsG husband.MsG and baby.MsG 'She leaves far looking for food for her husband and baby.' (FrWAC, harunyahya.fr)

We suppose that this difference can be accounted for by a tendency to avoid ambiguity. For human nouns, the singular D would favor a joint reading, which explains why it is quite rare with a split reading in the corpus. However, for non-human nouns, both the singular and plural D are possible with a split reading, and the singular D is more frequent, since non-human binomials are not ambiguous (or rarely ambiguous) between a joint and a split reading.

3.4 Animacy and Syntactic Functions

The syntactic position of the binomial may play a role as well. Since split binomials yield plural verb agreement (15), this might also favor plural D agreement. In (14), the split binomial is a complement and has a singular D. Similarly for animates, in (19a), the split binomial

| | subj | inv_ subj | comps | total |
|-------|------|-----------|-------|-------|
| Dpl | 42 | 17 | 503 | 562 |
| Dsg | 6 | 3 | 411 | 420 |
| total | 48 | 20 | 914 | 982 |

Table 4Agreement and syntactic function of two singular binomials inFrWAC

is a subject with a plural D, and in (19b) it is a complement with a singular D.

Animates tend to be more often in subject position than inanimates (Clark & Begun 1971). Previous studies have examples of (animate) binomials in subject position only (e.g., see (7b) and Le Bruyn & de Swart 2014 and Heycock & Zamparelli 2005). So it may be the independent tendency for animates to be in subject position that favors the plural D.

Since human nouns have a strong preference for plural D, and the singular D is quite rare as illustrated in table 3, it is difficult to compare the interaction between number agreement and syntactic position for human nouns. Thus, we only compare D number agreement for inanimate nouns in different syntactic positions in FrWac. In order to balance the singular/plural D, we chose two pairs of inanimate nouns which appear with both singular and plural D: *D nom et prénom* ('D name and first name') (324 tokens of Dsg and 513 tokens of Dpl), *D date et heure* ('D date and hour') (96 tokens of Dsg and 49 tokens of Dpl) and manually annotated their syntactic position: (preverbal) subject, inverted-subject, complement.

Table 4 shows that there are significantly more examples in complement position, and the plural D is preferred in subject position. But in complement position, the difference between singular/plural D is not significant.

Our results are consistent with the claim that conjoined human nouns favor plural agreement crosslinguistically (Corbett 1991). However, there are not enough examples of human nouns for a statistical study, and other semantic or pragmatic factors, such as topicality, definiteness, distinctness, may play a role too. This is why we decided to run two acceptability rating experiments to test the interaction of animacy and agreement carefully controlling agreement, and syntactic functions.

4 Two Acceptability Rating Experiments

In this section, we report two quantitative gradient acceptability judgment studies for D agreement in French singular split binomials. The first study focuses on a set of sentences with the binomials in subject position, with a plural verb, with half of the examples human and the other half, non-human. It aims to test the role of animacy, excluding other potential linguistic factors in a context forcing the split reading.

The second study is designed to test the role of syntactic function, and to test acceptability of binomials without plural priming. As in the first study, the binomials in subject position involve plural verb agreement, whether the verb is singular or plural may have an effect on the NP-internal agreement. In Experiment II, the experimental items are objects of a singular verbs, without any other plural agreement. In comparing these two studies, we can test the effects of this syntactic function bias. Since the object context is neutral between a split and a joint reading, we chose human and non-human nouns which cannot have a joint interpretation.

4.1 Experiment I 4.1.1 Materials

We constructed two lists of 12 experimental items, one with 12 human singular binomials (20a) and the other with 12 non-human singular binomials (20b), inspired from the corpus FrWac. The experimental items are in subject position and the verb is collective and in the plural form in order to force a split reading. The D is either singular or plural.

- **D**[-human]: Il arrive (20)a. souvent que it happen.3sG often that identifiant vos/votre et mot de passe ne your.pl/your.sg username.msg and password.msg neg pas reconnus par le site. soient be.subjv.3PL NEG recognized.PL by the site 'It often happens that your username and password are not recognized by the site.'
 - D[+human]: Les/Le b. directeur et the.pl/the.MSG director.MSG and du sous-directeur secteur se sont assistant director.MSG of.MSG sector.MSG REFL AUX.3PL d' accord sur le projet. mis put.M of agreement on the project 'The director and assistant director of the sector agreed on the project.'

We also included six control items without coordination. For each item, one version appears in a grammatical item (21a) and the other in an ungrammatical item (with number agreement error) (21b).

- (21) a. **Control1:** La tête dans les genoux, je the.FSG head.FSG in the.MPL knee.PL, I dormirais peut-être deux heures. sleep.COND.1SG perhaps two hour.PL
 - b. **Control2:** La tête dans le genoux, je the.FSG head.FSG in the.MSG knee.PL, I dormirais peut-être deux heures. sleep.COND.1SG perhaps two hour.PL

We also included a set of 15 items from another unrelated experiment as fillers. Each item has three conditions and there are 45 sentences in total.

4.1.2 Procedure

43 subjects participated in the experiment, recruited from the website RISC (http://www.risc.cnrs.fr/). Participants have to fill a survey before starting the experiments, reporting mainly their mother language, age, sex, etc. One participant was removed as he was not a native speaker and 42 were retained. Participants were asked to rate the acceptability of each sentence, with a Likert scale from 1 (completely unacceptable) to 10 (perfectly acceptable), which is the usual scale in the French school system.

The items are generated using a Latin square. Participants could only see one possible D (singular/plural) for each item, the number of which was counterbalanced across participants. The order of experimental items is also randomized in each trial.

4.1.3 Results

The results (figure 1) show that plural D are more acceptable than singular D, for both human and non-human nouns. Non-human binomials with a plural D (mean: 7.87) are more acceptable than the grammatical control items (mean: 7.3). The singular D (mean: 6.96) is less acceptable, but better than the ungrammatical control sentences (mean: 6.0). For human binomials, a plural D (mean: 7.13) is also judged better than a singular D (mean: 6.39).

We ran a maximum mixed-effects linear model with items and participants included as random factors (Barr et al. 2013), using *lmer4* package in R. The dependent variable is the mean of participants' ratings, and independent predictors are Noun animacy and Dsg/Dpl. There are significant main effects of D number (p = 0.03) and N animacy (p = 0.01): non-human nouns are more acceptable than the human nouns. And plural D are significantly more acceptable than singular ones. We didn't detect any effect of participants' age or so-cial gender.

4.1.4 Discussion

The results of Experiment I are consistent with our corpus data in the sense that inanimate (split) binomials are more acceptable than



Figure 1 Acceptability ratings of subject binomials with Dsg/Dpl

animate (split) binomials.

Furthermore, the plural D is preferred for human nouns, which is in line with our hypothesis because the singular D can be ambiguous between a split and joint reading. However, contrary to the corpus data, the plural D is also preferred for inanimate nouns in this experiment. We suppose that this may come from a semantic bias: we use plural verbs to force a split reading, so the plural verb agreement may have an effect on the preference for a plural D. We thus ran a second experiment with the binomials in object position.

4.2 Experiment II

We used the same experimental items as Experiment I but in object position. It is not possible to use the same context as Experiment I, but we tried to keep the contexts similar. We took the same pairs of nouns, 12 animate and 12 inanimate.⁵

⁵In the glosses, REF.1SG represents the reflexive pronoun of first person singular.

D[-human]: Chaque fois que je me (22)connecte, je a. time that I REF.1SG log-in, each dois taper mes/mon identifiant et must type my.PL/my.MSG username.MSG and mot de passe. password.msg 'Every time I log in, I need to type in my username and password' D[+human]: Il faudrait pouvoir prévenir b. it should be-able inform les/le directeur et sous-directeur the.PL/the.MSG director.MSG and assistant director.MSG de l'établissement. of the establishment 'It is necessary to be able to inform the director and assistant director of the establishment.'

The control items are the same as in Experiment I. We also included 24 items, for a total of 48 sentences as fillers.

4.2.1 Results

The procedure is the same as in Experiment I. We had 51 participants recruited from the RISC, all of them native speakers of French.

Figure 2 reports the mean for each condition of Experiment II. For non human nouns, both plural D (mean: 6.99) and singular D (mean: 6.83) are acceptable, but a little lower compared to the grammatical sentences (mean: 8.09). Human nouns are less acceptable than nonhuman nouns, both for plural D (mean: 5.97) and singular D (mean: 5.47), and singular D is even worse than the ungrammatical items (mean: 5.49).

Table 6 reports the results of the maximum mixed-effects linear model with the participants' rating as the dependent variable and D number and N animacy as independent variables. Only animacy has a significant effect (p < 0.001): non-human nouns are more acceptable than human nouns. However, we don't detect a significant effect



Figure 2 Acceptability ratings of object binomials with Dsg/Dpl

of D number (p = 0.19): there are no significant differences between singular and plural D in object position.

In object position, the preference for plural D doesn't exist anymore. Consistent with the first experiment, non-human nouns are still more acceptable than human nouns. However, contrary to Corbett (1991), we didn't detect any interaction between D agreement and N animacy, which means that the plural and singular D are judged in a similar way for both human and non-human nouns, something for which we are not able give an explanation in this paper.

5 General Discussion

5.1 Corpus Data and Experimental Data

We did a statistical comparison between Experiments I and II using a maximum mixed-effects linear model. The fixed effects consist of three factors: animacy, syntactic function and D number, as well as the interaction among them. For random effects, we included D



Figure 3 ggplot of the maximum linear model comparing Experiments I and II

number and noun animacy and their interaction for subjects, and the syntactic function of N and D number as well as their interaction for items.

As showed in figure 3, there are significant effects for D number (p = 0.03), animacy (p < 0.001) and syntactic function (p < 0.001). The interaction between animacy and syntactic function is also significant (p = 0.04).

We conclude that the acceptability of singular binomials is not a categorical phenomenon, but sensitive to several factors. In what follows, we would like to offer two considerations. First, the effects of animacy and syntactic position can be explained in terms of language processing difficulties caused by a tendency to avoid potential ambiguity. As a result, the difference of number agreement preferences between the corpus data and the experimental results may come from the bias that animate binomials are more frequent in subject position.

5.1.1 Semantic Ambiguity

The corpus data (section 3.2) have shown the different preferences for joint/split reading regarding animacy. Split is a higher, frequent dominant reading with non-human nouns, while for human nouns, it is a lower-frequency subordinate reading. When the second conjunct of the coordination appears, which was always congruent with the split reading, this will give rise to processing difficulties especially for the human nouns. Thus, this incremental processing difficulty may lead to lower judgments of acceptability.

Furthermore, recent models of ambiguity resolution (MacDonald et al. 1994, Tabor et al. 1997, McClelland et al. 1989) agree that readers are able to make use of available contextual information to help them activate and integrate the appropriate meaning of an ambiguous word or an ambiguous syntactic structure. In Experiment I, we use plural verbs to force the binomials in the subject position to have a split reading. The result illustrates that it is more acceptable with a context biased toward split reading (subject position with a plural verb) than with a neutral context (object position after a singular verb). We then explain the effects of syntactic function by its bias for the split reading: the binomials are more acceptable when the contextual information provides information that facilitates resolving the joint and split reading ambiguity.

We also detect a significant interaction between animacy and syntactic function. This means that the differences between subject and object position are greater for human nouns than for non-human nouns. As shown in figure 3, there are important differences between subject and object position for human nouns, while the differences are quite tiny for non-human nouns, especially for the singular determiner. We can interpret this effect by saying that the acceptability of human binomials depends more on the disambiguated context.

5.1.2 Agreement and Animacy

The empirical data revealed that the determiner can agree in number with the whole coordination (plural), or with the closest conjunct (singular). The preference for singular or plural D varies according to the syntactic position. Experiment I shows that a plural D is more acceptable with either human or non-human nouns in the subject position, while there is no preference for singular or plural D in object position in Experiment II.

The experimental data (section 4) are consistent with the corpus data (section 3) in that for non-human nouns, the plural D is preferred in subject position, while in object position, there is no difference between singular and plural D.

5.2 French Data and Previous Analyses 5.2.1 Optimality Theory

Le Bruyn & de Swart (2014) develop an analysis that supposes a different syntactic structure depending on the meaning. Using Optimality Theory (OT), they propose a ranking of constraints (23a), based on the typology of de Swart & Zwarts (2008) and de Swart & Zwarts (2009). The ranking of a specific constraint for coordination (*FunctSCoordP) with respect to a more general markedness constraint (*FunctN) and three faithfulness constraints ((23b)–(23d)) determine the article use with split and joint coordination readings.

- (23) a. {FPl, *FunctSCoordP} >> {FDr, FDef} >> *FunctN
 - b. FPI: reference to a plurality of individuals must be reflected in the form
 - c. FDr: the presence of a discourse referent in the semantics corresponds to an expression that carries discourse referential force
 - d. FDef: reference to discourse unique individuals requires the use of an expression of definiteness

*FunctN is a general constraint to avoid functional structure (D) in the nominal domain, of which *FunctNSCoordP is a specific subconstraint to avoid functional structure (D) on top of a split coordination, thus yielding *Man and wife were happy together*, in languages like English which do not allow bare nominal arguments (with singular count nouns). Since *FunctNSCoordP is ranked above *FunctN, coordination is an exception to the general rule enforced by the faithfulness constraints (requiring an explicit marking of definiteness). Since it is ranked above the faithfulness constraints FDr and FDef, split coordination relaxes the usual requirements on definite and indefinite articles.

The ranking of constraints may vary crosslinguistically. In languages like English or French, *FunctN is ranked low with respect to the faithfulness constraints governing article use and *FunctNSCoordP is ranked high, thus Split-Coordination Phrases behave differently from Joint-Coordination Phrases. Le Bruyn & de Swart (2014) proposed that for the joint reading, there is a DP projection above the coordination, and for the split reading (24a), the first noun is combined with the D to form a DP, which is combined with the second noun (24b).

(24) a. $[_{CoordP}[_{DP} D [NP and_{joint NP} NP]]$ b. $[_{CoordP}[_{DP} D NP] and_{split} [_{NP} NP]]$

The analysis of Le Bruyn & de Swart 2014 relies on linguistic intuitions and is based on the assumption that only one option is posible in a given language, which is unable to account for our results. Our empirical results suggest that number agreement for the joint reading is categorical, since only a singular D is allowed. However, both singular and plural D are allowed for the split reading, and which one is used is rather a matter of preference (cf. Gries 2003, Bresnan & Nikitina 2009, Bresnan & Hay 2008). Multiple factors, like semantic interpretation, animacy and the syntactic function of the binominal, play a role for computing agreement.

These constraints would probably be best represented with a competition algorithm (cf. MacWhinney et al. 1984, MacWhinney & Bates 1989, MacDonald et al. 1994). We rely on previous work within LFG (see Bresnan 1982, Dalrymple & Hristov 2010 and Kathol 1997 for an overview) and HPSG (Pollard & Sag 1994) to sketch an HPSG analysis.

5.2.2 An HPSG Analysis

An & Abeillé (2017) have proposed a simplified HPSG analysis for joint and split binomials. We use a general schema for coordina-



bare-joint-coord-phr bare-split-coord-phr NP-joint-coord-phr NP-split-coord-phr

Figure 4 Hierarchy of nominal coordinate phrases

tion, which is n-ary phrase (Borsley 2005, Abeillé 2005). The conjunction forms a subconstituent with the following conjunct, but it is not the head of the coordination, which is a subtype of *unheaded phrase*. We use a CONJ feature to distinguish conjuncts with a conjunction, and conjuncts without. *nelist* stands for non-empty-list, and DTRS (Daughters) for immediate constituents.

(25)
$$coord-phrase \Rightarrow \left[DTRS \ \left\langle nelist \left(\left[CONJ \ nil \right] \right) \oplus nelist \left(\left[CONJ \neq nil \right] \right) \right\rangle \right]$$

The syntactic structure is the same for split and joint coordination. We propose a cross-classification that distinguishes interpretation (split or joint) and lexicality (bare or NP coordination). This leads to the hierarchy in figure 4.

As proposed in Kathol 1997 and Wechsler & Zlatić 2003, we distinguish CONCORD and INDEX features. As illustrated in example (26) (repeated from (3a)), the verb shows INDEX agreement (plural) while the determiner shows CONCORD agreement (singular).

(26) This [boy and girl] are eating a pizza.

For *joint-coord-phr*, the schema imposes sharing of CONCORD and INDEX features:

 $(27) \quad joint-coord-phr \Rightarrow \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \text{CONCORD} & 1 \\ \text{INDEX} & i \end{bmatrix} \\ \text{DTRS} & \left\langle \begin{bmatrix} \text{CONCORD} & 1 \\ \text{INDEX} & i \end{bmatrix} \cdots \begin{bmatrix} \text{CONCORD} & 1 \\ \text{INDEX} & i \end{bmatrix} \right\rangle \end{bmatrix}$

This is compatible with bare nouns (28a) and NP coordination (28b).

- (28) a. Un [collègue et ami] est venu hier. a.MsG colleague and friend AUX. come.PST yesterday 'One colleague and friend come yesterday.'
 - b. [Un grand poète et un grand homme] a.MSG great.MSG poet.SG and a.MSG great.MSG man est mort hier.
 - is dead yesterday

'A great poet and a great man is dead yesterday.'

For *split-coord-phr*, the schema is (29) and the INDEX features are not shared:

(29)
$$split-coord-phr \Rightarrow \begin{bmatrix} CONT & [INDEX & i \oplus ... \oplus n] \\ DTRS & \langle [INDEX & i] ... & [INDEX & n] \rangle \end{bmatrix}$$

For bare split nominals, the schema in figure 5 allows both plural and Closest Conjunct Agreement. Following Villavicencio et al. (2005), we use RAGR for right-agreement (rightmost noun, N2) and LAGR for left-agreement (leftmost noun, N1). We consider the nouns to be the syntactic head, and use the SPR feature (Specifier) for D selection. We ignore gender features here, for which Abeillé et al. (2018) show that only Closest Conjunct Agreement is allowed (recall (14)).

The schema bare-split-coord-phr is illustrated in figure 6 for votre/vos



Figure 5 Schema for *bare-split-coord-phr*

nom et prénom. Notice that the closest conjunct (the leftmost noun) is also the highest.

However, this analysis also has its limits. It accounts for the categorical difference between joint and split reading, and for the different agreement possibilities, but it does not take into account the preferences for the singular/plural D in the split case.

6 Conclusion

Previous studies have addressed number agreement of binominal coordinations in different languages (e.g., Le Bruyn & de Swart 2014, King & Dalrymple 2004, Demonte & Perez-Jimenez 2012, Villavicencio et al. 2005). We use large-corpus data and experimental results to investigate this question. Furthermore, we explored semantic and syntactic factors that may affect number agreement of the shared determiner.

Our empirical findings are twofold. First, we show that Closest Conjunct Agreement is attested in French and is not superficial since it is sensitive to syntactic (word order and syntactic function) and semantic factors (noun animacy). Since in D-Noun agreement the target is before the controller, the possibility of Closest Conjunct Agreement is expected by Corbett's agreement hierarchy, especially for non-human nouns. As the closest noun is also the highest one,



Figure 6 An illustration of *bare-split-coord-phr*

it is also expected under a structural account of agreement (Marušič et al. 2007). However, a plural D is preferred in subject position for both human and non-human nouns, contrary to other Romance languages, but no differences are observed in the object position.

Second, the binomial is ambiguous between a joint or a split reading, and human nouns favor the joint reading whereas non-human nouns favor the split reading. In the split context, human nouns are less acceptable than non-human nouns. Furthermore, they are more acceptable in a disambiguated context, for instance, in subject position.

Third, we argue that a traditional approach using uncontrolled grammatical judgments can lead to empirical inadequacy as in the OT approach of Le Bruyn & de Swart (2014).

Finally, new questions arise: whether Closest Conjunct Agreement is also possible in other types of agreement in French, such as nounadjective agreement and subject-verb agreement, and whether it is sensitive to the same factors.

Acknowledgments We thank the editor, Christopher Pinon, as well as the anonymous reviewer for their feedback on previous versions of this paper. This work is supported by strand 2 of LabEx EFL(ANR-10-LABX-0083) and China Scholarship Council. Special thanks to Olivier Bonami, Berthold Crysmann, Jean-Pierre Koenig, Jeremy Kuhn and Bert Lebruyn for help-ful comments and discussions. We are also grateful to the audience of the CSSP 2017 conference.

References

- Abeillé, A. (2005). Les syntagmes conjoints et leurs fonctions syntaxiques. *Langages*, *160*(4), 42–66.
- Abeillé, A., An, A., & Shiraïshi, A. (2018). L'accord de proximité du déterminant en français. *Discours*, 22.

URL http://journals.openedition.org/discours/9542

An, A., & Abeillé, A. (2017). Agreement and interpretation of binominals in French. In S. Müller (Ed.) *Proceedings of the 24th International Conference on Head-Driven Phrase Structure Grammar, University of Kentucky,* Lexington, (pp. 26-43). Stanford, CA: CSLI Publications.

URL http://cslipublications.stanford.edu/HPSG/2017/ hpsg2017-an-abeille.pdf

- Baroni, M., Bernardini, S., Ferraresi, A., & Zanchetta, E. (2009). The WaCky wide web: A collection of very large linguistically processed webcrawled corpora. *LREC Conference*, 43(3), 209–226.
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3), 255–278.
- Borsley, R. D. (2005). Against ConjP. Lingua, 115(4), 461-482.
- Bresnan, J. (Ed.) (1982). *The Mental Representation of Grammatical Relations*. The MIT Press.
- Bresnan, J., & Hay, J. (2008). Gradient grammar: An effect of animacy on the syntax of *give* in New Zealand and American English. *Lingua*, 118(2), 245–259.
- Bresnan, J., & Nikitina, T. (2009). The gradience of the dative alternation. In L. U. . L. H. Wee (Ed.) *Reality Exploration and Discovery: Pattern Interaction in Language and Life*, (pp. 161–184). Stanford, CA: CSLI.
- Clark, H. H., & Begun, J. S. (1971). The semantics of sentence subjects. *Language and Speech*, 14(1), 34–46.
- Comrie, B. (1989). *Language Universals and Linguistic Typology: Syntax and Morphology*. The University of Chicago Press.
- Corbett (1991). Gender. Cambridge University Press.
- Croft, W. (2002). Typology and Universals. Cambridge University Press.
- Dalrymple, M., & Hristov, B. (2010). Agreement patterns and coordination in Lexical Functional Grammar. In M. Butt, & T. H. King (Eds.) *Proceedings of LFG10*, (pp. 187–207).
- de Swart, H., & Zwarts, J. (2008). Article use across languages: An OT typology. In A. Grønn (Ed.) *Proceedings of Sinn und Bedeutung* 12, (pp. 628–644). University of Oslo.
- de Swart, H., & Zwarts, J. (2009). Optimization principles in the typology of number and articles. In B. Heine, & H. Narrog (Eds.) *The Oxford Handbook of Linguistic Analysis*, (pp. 555–581). Oxford University Press.

Demonte, V., & Perez-Jimenez, I. (2012). Closest conjunct agreement in Spanish DPs: Syntax and beyond. *Folia Linguistica*, 46(1), 21–73.

Gries, S. T. (2003). Towards a corpus-based identification of prototypical

instances of constructions. *Annual Review of Cognitive Linguistics*, (pp. 1–27).

- Haspelmath, M. (2013). Occurrence of nominal plurality. In M. S. Dryer, & M. Haspelmath (Eds.) *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. URL https://wals.info/chapter/34
- Heycock, C., & Zamparelli, R. (2005). Friends and colleagues: Plurality, coordination, and the structure of DP. *Natural Language Semantics*, 13(3), 201–270.
- Kathol, A. (1997). Agreement and the syntax-morphology interface in HPSG. In R. Levine, & G. Green (Eds.) *Studies in Contemporary Phrase Structure Grammar*, (pp. 223–274). Cambridge University Press.
- Kayne, R. S. (1994). The Antisymmetry of Syntax. The MIT Press.
- King, T. H., & Dalrymple, M. (2004). Determiner agreement and noun conjunction. *Journal of Linguistics*, 40(1), 69–104.
- Le Bruyn, B., & de Swart, H. (2014). Bare coordination: The semantic shift. *Natural Language and Linguistic Theory*, 32(4), 1205–1246.
- MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review*, 101(4), 676–703.
- MacWhinney, B., & Bates, E. (1989). *The Crosslinguistic Study of Sentence Processing*. Cambridge University Press.
- MacWhinney, B., Bates, E., & Kliegl, R. (1984). Cue validity and sentence interpretation in English, German, and Italian. *Journal of Verbal Learning and Verbal Behavior*, 23(2), 127–150.
- Marušič, F., Nevins, A., & Saksida, A. (2007). Last-conjunct agreement in Slovenian. In R. Compton, M. Goledzinowska, & U. Savchenko (Eds.) *Formal Approaches to Slavic Linguistics*, vol. 15, (pp. 210–227). Ann Arbor: Michigan Slavic Publications.
- McClelland, J. L., St. John, M., & Taraban, R. (1989). Sentence comprehension: A parallel distributed processing approach. *Language and Cognitive Processes*, 4(3-4), 287–335.
- Murphy, A., & Puškar, Z. (2018). Closest conjunct agreement is an illusion. *Natural Language and Linguistic Theory*, *36*(4), 1207–1261.
- Namer, F. (2000). FLEMM: un analyseur flexionnel du français à base de règles. *Traitement automatique des langues*, *41*(2), 523–547.

- Pollard, C., & Sag, I. A. (1994). *Head-Driven Phrase Structure Grammar*. The University of Chicago Press.
- Tabor, W., Juliano, C., & Tanenhaus, M. K. (1997). Parsing in a dynamical system: An attractor-based account of the interaction of lexical and structural constraints in sentence processing. *Language and Cognitive Processes*, 12(2-3), 211–271.
- Villavicencio, A., Sadler, L., & Arnold, D. (2005). An HPSG account of closest conjunct agreement in NP coordination in Portuguese. In S. Müller (Ed.) Proceedings of the 12th International Conference on Head-Driven Phrase Structure Grammar, Department of Informatics, University of Lisbon, (pp. 427–447). Stanford, CA: CSLI Publications.
- Wechsler, S., & Zlatić, L. (2003). *The Many Faces of Agreement*. Stanford, CA: CSLI Publications.
- Zaenen, A., Carletta, J., Garretson, G., Bresnan, J., Koontz-Garboden, A., Nikitina, T., O'Connor, M. C., & Wasow, T. (2004). Animacy encoding in English: Why and how. In *Proceedings of the 2004 ACL Workshop on Discourse Annotation*, (pp. 118–125). Association for Computational Linguistics.

Extreme Adjectives in Comparative Structures and *even*

Micky Daniels • Yael Greenberg

Abstract This paper examines two related puzzles, observed in the literature about extreme adjectives (see, e.g., Paradis 2001, Rett 2008, Morzycki 2012), namely, (a) why such adjectives are questionable within comparative structures and (b) why and how exactly the presence of even improves the felicity of such constructions. After examining the solutions proposed in Morzycki 2012 for these two puzzles, we propose an alternative solution which integrates three components: (i) the fact that extreme-adjective comparatives necessarily presuppose the "positive form" of these adjectives (building on Morzycki's semantics for extreme adjectives); (ii) an updated, gradability-based semantics for even (Greenberg 2015, 2018), which guarantees that comparatives with even presuppose the corresponding "positive form," with all kinds of adjectives (extreme and non-extreme alike); and (iii) a local Maximize Presupposition!-type effect, such as that suggested by Singh (2011), leading to the preference of the extreme-adjective comparative with even over a competing alternative without it. While the latter component presents a number of challenges requiring further research, the proposal is shown to be supported by cross-linguistic data and by comparing extreme and lower-closed adjectives in terms of scale structure and behavior in comparatives.

Keywords extreme adjectives · comparatives · degrees · standards · *even* · scalar presupposition · Maximize Presupposition! · scale structure

Y. Greenberg, Bar-Ilan University, yaelgree@gmail.com

M. Daniels, Bar-Ilan University, mickydaniels@gmail.com

In Christopher Pinon (Ed.), Empirical Issues in Syntax and Semantics 12, 61–89. Paris: CSSP. http://www.cssp.cnrs.fr/eiss12/ © 2019 Micky Daniels & Yael Greenberg

1 Introduction

This paper discusses two puzzling observations which have been noted in the literature on extreme adjectives (such as *gigantic, ginormous, excellent, scrumptious, terrible*) in comparative constructions (e.g., Paradis 2001, Rett 2008, Morzycki 2012). The first puzzle is why such adjectives are usually judged to be degraded within the comparative, as in (1b), as opposed to the case with non-extreme adjectives, as in (1a). The second puzzle is why the felicity of extremeadjective comparatives (henceforth EA comparatives) is improved with even, as seen in (1c).

- (1) a. Godzilla is bigger than Mothra.
 - b. ??Godzilla is more gigantic than Mothra. (Morzycki 2012:(10a))
 - c. Godzilla is even more gigantic than Mothra. (Morzycki 2012:(61b))

In this paper we suggest a solution for these two puzzles and examine some implications of this solution for issues such as the semantics of extreme adjectives, their associated scale structure, comparatives and gradability, as well as the semantics of even. We will start in section 2 with a review of previous solutions to these two puzzles, proposed by Morzycki (2012), and in section 3 raise a few challenges to these solutions. In section 4, we will present our proposal for solving both mysteries, which integrates three building blocks: (i) Morzycki's (2012) semantics of EA comparatives, (ii) Greenberg's (2015, 2018) "gradability-based" semantics for the scalar presupposition of even and (iii) a principle akin to localized Maximize Presupposition!, such as that proposed in Singh 2011. We will provide support for our proposal based on some cross-linguistic data in section 5.1, and by examining some differences between extreme adjectives and lower-closed ones in section 5.2. Finally, in section 6, we summarize and point out some open issues and directions for future research.

2 Background

2.1 Extreme Adjectives in the Comparative

Morzycki (2012) discusses several ways to identify extreme adjectives. Among these ways, one is that such adjectives (e.g., *enormous*) occur felicitously with a specific group of degree modifiers, for instance, *absolutely*, *full-on*, *downright*, *flat-out*, *positively*, as in (3), but not with, for example, *very*, as in (2). In contrast, non-extreme adjectives, such as *big*, are infelicitous with *downright* (see (3)), but felicitous with *very* (see (2)).

- (2) very ??enormous / big
- (3) downright enormous / ??big (Morzycki 2012:(5,4))

Another feature of extreme adjectives, noted by, among others, Paradis (2001) and Morzycki (2012), is that, as mentioned above, they are degraded in comparative structures (as in (1b), (4b), (5b) and (6b)), unlike their non-extreme counterparts (as in (1a), (4a), (5a) and (6a)):

- (4) a. A is better than B.b. ??A is more excellent than B. (Paradis 1997)
- (5) a. The salsa is worse than the guacamole.b. ??The salsa is more terrible than the guacamole.

(Portner & Rubinstein 2016:(21))

(6) a. Jane is more beautiful than Dorothy.b. ??Jane is more gorgeous than Dorothy.

To capture this latter observation, Morzycki relies on his analysis of extreme adjectives, which is based on the idea that such adjectives are associated with degrees which are "off the scale," that is, exceed the salient part of the scale. More specifically, Morzycki argues that in any given context, where gradable adjectives are associated with a scale, our attention is not on the entire scale but on a salient portion of it. This leads Morzycki to add a component to the semantics of non-extreme adjectives, where besides denoting relations between individuals and degrees, type $\langle e, \langle d, t \rangle \rangle$ (as suggested in, e.g., Kennedy & McNally 2005) these degrees are within the contextually salient portion of the scale *C*, as seen in (7) for *big*:

(7)
$$\llbracket \operatorname{big}_C \rrbracket = \lambda x \lambda d. d \in C \wedge \operatorname{big}(d)(x)$$

According to this, big_C is a function from an entity x and a degree d, that returns truth iff d is a member of the salient portion of the scale C (the portion of the scale associated with big), and x is big to degree d.

Morzycki then suggests that extreme adjectives are similar to their non-extreme counterparts in denoting relations between individuals and degrees. However, crucially, in the case of extreme adjectives, these degrees lie beyond the contextually provided scale. For example, $gigantic_C$ is interpreted as in (8), where *d* exceeds the maximum degree on the salient portion of the bigness scale *C*:

(8)
$$\llbracket \text{gigantic}_C \rrbracket = \lambda x \lambda d.d > \max(C) \land \operatorname{big}(d)(x)$$

Then, adopting a semantics for comparatives as in Kennedy 2004, Morzycki proposes that the EA comparative in (1b) would have the semantics in (9):

(9) $[[more gigantic_C than Mothra (is gigantic_C)]] = \lambda x. \exists d'(d' > max(C) \land big(d')(x) \land d' > d \land d > max(C) \land big(d)(Mothra))$

In prose, *more gigantic than Mothra* is true of an individual x iff there is a degree d' to which x is big which exceeds the salient portion of the scale C and this degree is higher than the maximal degree d to which Mothra is big, which also exceeds the maximal degree in C.

2.2 Morzycki's (2012) Explanations for the Two Puzzles

2.2.1 Morzycki's Explanation for the First Puzzle: Why are EA Comparatives Degraded?

Morzycki (2012) proposes two explanations for this puzzle. First, he suggests that EA comparatives are degraded since the act of comparing inherently makes degrees salient, thus leading to a pragmatic clash when applied to the non-salient degrees associated with extreme adjectives. Portner & Rubinstein (2016:15) give a more intuitive characterization of this clash in the case of (5b) (*??The salsa is more terrible than the guacamole*), which is infelicitous since

if the salsa is terrible, it is so overwhelmingly bad that it might be difficult or pointless to decide whether it is better or worse than the (also terrible) guacamole. After all, if it's terrible, you know all you need to know: that you're not going to eat it.

Another potential reason for this infelicity, suggested by Morzycki, is related to the maximality function, shown in the semantics of the comparative in (9). This function triggers an existential presupposition, that is, presupposes that there is a degree on the "giganticness scale," that is, a degree beyond the salient degrees in *C*, to which Mothra is big. Consequently, in (9) it is presupposed that Mothra is gigantic. Morzycki shows that this inference is, indeed, presupposed, as it survives negation, as in (10):

(10) ??Godzilla is (not) more gigantic than Mothra.
 Presupposes: Mothra is gigantic.
 (Adapted from Morzycki 2012:(37a))

The existence of this presupposition in EA comparatives leads Morzycki to propose that their infelicity is caused by difficulty in accommodating this presupposition. For example, in (1b), the difficulty would be to accommodate 'Mothra is gigantic'.

2.2.2 Morzycki's Explanation for the Second Puzzle: Why does even Improve the Felicity of EA comparatives?

Morzycki appears to base his explanation for this puzzle on the longenduring traditional semantics of *even*, according to which *even* ppresupposes that p is less likely than any other relevant focus alternatives q (cf. Horn 1969, Karttunen & Peters 1979, Rooth 1985, 1992, Guerzoni 2003, Chierchia 2013). He then proposes (p. 25) that

because *even* (...) is reflecting what is more or less expected in the discourse (Rooth 1985, Wilkinson 1996, Rullmann 1997, Giannakidou 2007), it provides a way for the speaker to acknowledge that the intended comparison is beyond the expected range, and to invite other discourse participants to play along.

According to Morzycki, then, this is why the presence of *even* helps language users overcome the difficulty in accommodating the presupposition that the degree of the compared elements is "beyond the expected range," or more technically, beyond the relevant standard.

3 Issues with Morzycki's Proposed Solutions

Regarding the proposed solution for the first puzzle, that is, the infelicity of EA comparatives, while Morzycki's idea that there is a pragmatic clash (between the non-salient degrees which are associated with extreme adjectives and the inherently salient degrees associated with compared entities) is intuitively a compelling reason for infelicity, it is still rather vague. As Portner & Rubinstein (2016) point out, without a clear definition of what salience of degrees amounts to, this explanation is rather difficult to test and evaluate.

As noted above, Morzycki's other possible explanation for the degraded status of EA comparatives is that there is a difficulty in accommodating the existential presupposition that is triggered in such comparatives, which leads to the inference that, for example, Mothra is gigantic. The problem with this solution, however, is that presupposition accommodation is quite common, and usually does not result in infelicity. It is unclear why there would be a special obstacle to accommodation specifically in EA comparatives.

As for the improved felicity of EA comparatives in the presence of *even*, as noted above, Morzycki appears to employ the "traditional" likelihood-based semantics for *even*. For this semantics to be insightful with respect to the improved felicity effect of *even* with EA comparatives, it would have to interact in some way with the possible reasons suggested for the original infelicity of such comparatives. However, none of the two suggested reasons for infelicity seem to be mitigated by this semantics of *even* in a sufficiently clear way.

Let us assume, for example, that in (1c) (*Godzilla is even more gigantic than Mothra*), *even* associates with *more*, so the prejacent of *even*, *p*, is *Godzilla is* [*more*]_F *gigantic than Mothra*. In this way, possible alternatives, *q*, could be *Godzilla is as gigantic as Mothra*, *Godzilla is less gigantic as Mothra*, etc. Crucially, presupposing that *Godzilla is more gigantic than Mothra* is less likely than such alternatives, *q*, does not affect in any clear manner the saliency of any of the degrees on the "size" scale involved in the semantics of the extreme adjective gi*gantic*. The "pragmatic clash" between the degrees made salient by the semantics of the comparative, and the non-salient degrees introduced by the semantics of the extreme adjective gigantic, does not seem to disappear, given the presence of *even*.¹

As for the second explanation for the oddness of EA comparatives suggested by Morzycki, namely, the difficulty in accommodating that the source of comparison has a degree on the extreme portion of the scale (e.g., that Mothra is gigantic in (10)), the likelihood-based scalar presupposition for *even* does not appear to address this reason for infelicity either. This is because assuming that p (*Godzilla is* [*more*]_F *gigantic than Mothra*) is less likely than q (e.g., *Godzilla is as gigantic*)

¹The same is true with other types of alternatives triggered by focused elements in such sentences: for example, where *p* is $[Godzilla]_F$ is more gigantic than Mothra, it would not solve the pragmatic clash or alter the salience of the associated degrees to assume that the Kraken is more gigantic than Mothra is a more likely alternative.

as Mothra) does not in any clear way facilitate accommodation of the presupposition that Mothra is gigantic. Such is the case for other focus placements and alternatives as well.

As an interim summary, it appears that assuming a likelihoodbased semantics of *even* is not fruitful here in that it does not seem to solve the suggested causes for infelicity of EA comparatives in any clear manner.

4 Our Proposal

As noted above, our solutions to the two puzzles integrate three main components, briefly reviewed in the next three subsections.

4.1 First Component: EA Comparatives Presuppose the "Positive Form" for both Source and Target

Our starting point is Morzycki's observed presuppositional pattern in (10), which appears to be correctly predicted by his proposed semantics for extreme adjectives (in (8)) and the presence of the maximality function in the comparative (in (9)). As noted above, in (1b) (??*Godzilla is more gigantic than Mothra*), for example, Morzycki takes this EA comparative to presuppose that Mothra is gigantic. Following Kennedy and McNally's (2005) terminology, we will assume that (1b) presupposes that Mothra is [pos] gigantic, that is, the degree to which the source of comparison, Mothra, is gigantic, is at least as high as the membership standard for giganticness (and the same is true for the target of comparison, Godzilla, which has a degree of giganticness which is higher still).

This move from the existential presupposition triggered by the maximality operator in the comparative to the "positive form" presupposition can, indeed, be derived from Morzycki's semantics of extreme adjectives. This is because according to his analysis, the scales associated with, for example, *big* and *gigantic*, are two subsections of one scale which lie "back to back." The smallest *gigantic* (extreme/non-salient) degree would be just above the greatest *big* (non-extreme/salient) degree, that is, just above *C*. Thus, the minimal degree above *C* would mark the location of the standard, above


Figure 1

which all degrees are extreme, as demonstrated in figure 1.

Since the maximality function in the comparative ensures that the source of comparison has a degree on the extreme subscale, that is, above *C*, and since Morzycki's characterization of extreme adjectives ensures that any degree above *C* is at least as high as the standard of giganticness, EA comparatives as in (1b) (*??Godzilla is more gigantic than Mothra*) indeed presuppose the "positive form" for both source and target (that Mothra (and Godzilla) is [pos] gigantic).

Note that this "positive form" presupposition in EA comparatives, has led, among others, Rett (2008) to claim that extreme adjectives have, in fact, lower-closed scale structure (based on Kennedy and McNally's (2005) scheme). This conclusion seems to be based on the apparent existence of a similar presuppositional pattern to that observed in extreme adjectives in lower-closed scalar adjectives, as in (11).

(11) This rag is (not) wetter than the chamois. *Presupposes:* The chamois is [pos] wet.

Despite this apparent similarity, we will show in section 5.2 that there is an important difference between the two types of comparatives, which will ultimately strengthen our proposal.

4.2 Second Component: An Updated Semantics for *even* with Intrinsic Sensitivity to Standards

We saw above that relying on the traditional, likelihood-based scalar presupposition of *even* is not fruitful for clearly understanding why its presence improves the felicity of EA comparatives. A theory of *even* that we find to be more helpful in this respect is Greenberg 2015, 2018. Greenberg points out several problems for the traditional likelihood-based account for *even*, which leads her to develop an updated, "gradability-based" scalar presupposition for this particle.

For our purposes, the most relevant component in this work is the claim that *even* includes reference to standards of comparison as an intrinsic part of its semantics. Specifically, Greenberg argues that a sentence with *even* presupposes that a non-focused element in both its prejacent, p, and its focus alternatives, q, must have a degree which is at least as high as the standard on a scale associated with a contextually supplied gradable property, G. An example adapted from Greenberg 2015 supporting this view is in (12).

(12) Context: John and Bill are players who applied to join our basketball team, where the standard of height is 1.90m. Their candidacy is being considered.

A: What about John and Bill? Should we recruit them?

- a. B: Well, John is 1.95m tall. Bill is (even) $[2.10]_F$. (We can recruit both.)
- B: Well, John is 1.70m tall. Bill is (??even) [1.75]_F.
 (We should not recruit either one.)
- c. B: Well, John is 1.75m tall. Bill is (??even) [1.95]_F. (We can recruit Bill.)

As Greenberg notes, where *even* is present, only (12a) is felicitous, as this is the only sentence where Bill's degree of height in p (2.1om) and in the alternative, q (1.95m), is in both cases at least as high as the relevant standard for height in the context (1.90m). The sensitivity of *even* to standards is further illustrated by its effect in comparatives,

as in (13).

(13) John is 1.70m tall. Bill is (even) taller (than that).(Greenberg 2018:(34))

As noted in Greenberg 2018, only the variant with *even* in (13) triggers a presupposition that 1.70m is at or above the threshold for tallness; that is, the presence of *even* leads to the presupposition of the "positive form," namely, that John, (and subsequently also Bill), is [pos] tall. Indeed, when *even* is absent, there is no problem uttering ... *but both are short* after (13), while crucially, with *even* this would lead to infelicity.

To capture this sensitivity of *even* to standards as well as other novel observations regarding *even*, Greenberg builds on an intuition in Rullmann 2007, according to which *even* indicates that p and its alternatives "are correlated with some graded property q." To formally capture this intuition, she relies on Beck's (1997) analysis of comparative correlatives and redefines the scalar presupposition of *even* as in (14), where x stands for a non-focused element within p, and G stands for a gradable property:

(14) even(C)(p)(w) is defined iff $\forall q \in C \land q \neq p \longrightarrow$ $\forall w_1, w_2(w_1 Rw \land w_2 Rw \land w_2 \in p \land w_1 \in q \land \neg p) \longrightarrow$ $\max(\lambda d_2.G(d_2)(x)(w_2)) > \max(\lambda d_1.G(d_1)(x)(w_1)) \land$ $\max(\lambda d_1.G(d_1)(x)(w_1)) \ge \operatorname{stand}_G$

In prose: *x* is more *G* in all accessible *p*-worlds than in all accessible [*q* and not *p*]-worlds and in the [*q* and not *p*] worlds, *x*'s degree of *G* is at least as high the standard for *G*. In the case of (12a), for example (*John is 1.95m tall. Bill is (even)* $[2.10]_F$), we can assume that we are measuring degrees to which Bill (a non-focused element in *p*) is tall (or alternatively, degrees to which he is suitable for the basketball team). Given (14), then, (12a) presupposes that (i) Bill's degree of tallness in all worlds where he is 2.10m tall is greater than in the worlds where he is 1.95m tall (and not 2.10m tall), and (ii) that in the

latter world set, John is considered to be [pos] tall, that is, his degree of tallness is at least as high as the contextually supplied standard. Since this presupposition is indeed met in (12a), the presence of *even* is felicitous, whereas in the parallel (12b-c), the second conjunct of the presupposition fails, so *even* is infelicitous.

As shown in Greenberg 2018, applying this scalar presupposition of *even* to the comparative in (13), would be as in (15), where p is *Bill is taller than John*, the alternative, q, is *Bill is as tall as John* and x (the non-focused element in p is Bill:

(15) $\forall w_1, w_2(w_1 Rw \land w_2 Rw \land w_2 \in (\exists d(\operatorname{tall}(d)(\operatorname{Bill}) \land d > 1.70m) \land w_1 \in [\exists d(\operatorname{tall}(d)(\operatorname{Bill}) \land d \ge 1.70m) \land \neg \exists d(\operatorname{tall}(d)(\operatorname{Bill}) \land d > 1.70m)] \longrightarrow \max(\lambda d_2.\operatorname{tall}(d_2)(\operatorname{Bill})(w_2)) > \max(\lambda d_1.\operatorname{tall}(d_1)(\operatorname{Bill})(w_1)) \land \max(\lambda d_1.\operatorname{tall}(d_1)(\operatorname{Bill})(w_1)) \ge \operatorname{stand}_{TALL})^2$

In prose: (i) Bill's degree of tallness in all accessible worlds where he is taller than 1.70m, is higher than in all worlds where he is exactly 1.70m tall (this is, of course, trivially met), and (ii) Bill's degree of tallness in the latter set of worlds is at least as high as the standard for tallness (i.e., he is tall). To illustrate, see figure 2 (based on Greenberg 2015).

We are now in a position to apply this analysis of comparatives with *even* to the EA comparative in (1c) (*Godzilla is even* [*more*]_F gigan*tic than Mothra*), where an alternative *q* is, for example, *Godzilla is as* gigantic as Mothra. Greenberg's scalar presupposition for *even* would now require that (i) Godzilla's size is greater on a scale of giganticness in the *p*-worlds (where Godzilla's size exceeds that of Mothra) than its size in the [*q* and not *p*]-worlds (where Godzilla's size equals that of Mothra) and crucially (ii) Godzilla's size in the latter set of worlds is at least as high as the standard for giganticness. To clarify,

²For further details on the "gradability-based" semantics of *even* and its application on comparative structures, beyond what was described here, see Greenberg 2015, 2018.





see figure 3.

What will be crucial in what follows is that given this analysis, the same "positive form" presupposition (i.e., that the source and target have a degree which is at least as high as the standard) is triggered by both EA comparatives, given Morzycki's (2012) semantics (as described in section 4.1 above), and by comparatives with *any* adjective in the presence of *even*, given Greenberg's (2015, 2018) "gradability-based" semantics of *even*.

Given these two occurrences of the same presupposition, one potential way to explain the improved felicity of EA comparatives with *even* may be an effect similar to Maximize Presupposition!, to which we now turn our attention.

4.3 Third Component: a Maximize Presupposition!-like Effect The idea behind Maximize Presupposition! (Heim 1991, Sauerland 2008, Percus 2006, Chemla 2008, Singh 2011, Schlenker 2012) is that given two competing alternative utterances, which bear the same assertive content, the variant with the stronger presupposition that is satisfied in the context will be favored. Thus, the utterance with the same assertive content, but which has a weaker or no presupposition, will be degraded. For this reason, for example, (16a) is taken to be rejected in favor of (16b):



Figure 3

(16) a. #A sun is shining.

b. The sun is shining. (Singh 2009:(1a,1b))

The uniqueness of the sun, which is taken to be true in the context, based on real-world knowledge, is presupposed by the definite article *the* in (16b), hence the sentence in (16b) is favored over (16a), which does not trigger this presupposition, and which is, thus, infelicitous.

In what immediately follows, we will describe some proposed deviations from the more classic analyses of Maximize Presupposition!, which will come into play when we apply this principle to the case of EA comparatives and *even*.

The classic accounts of Maximize Presupposition! such as Heim 1991, Chemla 2008 and Sauerland 2008, define it as a principle which operates globally, at the root. This assumption does not hinder Maximize Presupposition!'s ability to account for examples such as (16). However, Percus (2006), followed by Singh (2011) noted examples, which challenge this view. Consider (17a) and (17b).

(17) a. #If Mary has exactly two students, she assigned the same exercise to all of them. b. If Mary has exactly two students, she assigned the same exercise to both of them.
 (Adapted from Percus 2006:(30))

The challenge posed by (17) to Maximize Presupposition! taking effect globally, is that (17a) (with *all*) is rejected in favor of (17b) (with *both*), despite the fact that as a whole, neither sentence presupposes that Mary has exactly two students. It appears that the latter presupposition is satisfied thanks to the antecedent, and that (17b), where this same presupposition is triggered by *both*, is thus favored over (17a) with *all*, which does not.

To account for such sentences with Maximize Presupposition!, Percus (2006) suggests that it causes the preference of an alternative lexical item (in this case, *both*), which on its own triggers a stronger presupposition than the other alternative, regardless of the global presupposition of the sentence. Alternatively, Singh (2011) suggests that Maximize Presupposition! takes effect on the level of subclauses, evaluated in their "local context." Thereby in (17), what causes Maximize Presupposition! to take effect (and (17b) to be favored) is that in the consequent (*she assigned the same exercise to both* vs. *she assigned the same exercise to all*) there is a difference in presuppositional strength, and the presupposition of *both* is satisfied in the local context (i.e. the initial context updated with the antecedent).

A further example indicating a case which deviates from the classic characterization of Maximize Presupposition! is in (18):³

- (18) a. #All of the two students are nice.
 - b. Both of the two students are nice.

As opposed to the scenario in (17), where globally neither sentence presupposes that there are exactly two students, in (18), both sentences presuppose this, due to the presence of *of the two students*. Thus, it appears that (18b), which is presuppositionally stronger (be-

³Thanks to an anonymous EISS reviewer for this example and related insights, based on his/her personal communication with Amir Anvari.

cause of the presence of *both*), is favored due to the fact that another element in the sentence (namely, *the two students*) triggers the same presupposition as *both* does. We propose that this case can be explained by an effect similar to localized Maximize Presupposition! à la Singh 2011,⁴ in the following way: when *the two students* (which appears syntactically below *both*) is computed, it creates a local context in which its presupposition is assumed to be satisfied. This local context then causes the preference of (18b) with *both*, which triggers the same presupposition, over (18a), which does not.

A final digression from classic Maximize Presupposition!, which will be significant later on, is one proposed by Amsili & Beyssade (2006). The latter argue that Maximize Presupposition! will also take effect if the disfavored competing expression is a null form and not an overt form as in the default case. Consider (19a) and (19b).

| (19) | a. | #Jean est malade. Marie est malade \emptyset . | | |
|------|----|--|--|--|
| | | 'John is sick. Mary is sick \emptyset .' | | |
| | h | Ioan ost malado. Mario ost malado aus | | |

b. Jean est malade. Marie est malade aussi.
'John is sick. Mary is sick too.' (Amsili & Beyssade 2006:(11b))

Here, (19b) with the additive particle *aussi*, which triggers the existential presupposition that someone else in the context who is not Mary is sick as well, is favored over (19a) with the null form \emptyset , where this presupposition (which holds in the context) is not triggered.

4.4 Integrating the Three Components

Returning to our original example, see (20a) and (20b) (repeated and adapted from (1b) and (1c)), below:

- (20) a. ??Godzilla is \emptyset more gigantic than Mothra.
 - b. Godzilla is even more gigantic than Mothra.

We are now in a position to explain the contrast between the EA

⁴Percus's (2006) analysis appears to be equally applicable.

comparative in (20a) and the EA comparative with *even* in (20b), building on the three components described in the previous sections.

First, we established that EA comparatives presuppose the "positive form" for the source (and consequently also the target) of comparison. We did this by following Morzycki 2012, in claiming that the interaction of the maximality function in the comparative, triggering an existential presupposition over degrees, with the semantics of extreme adjectives, leads to a situation where the source of comparison is independently presupposed to have a degree which is at least as high as the standard related to the extreme adjective scale.

Second, we argued that *even*, too, presupposes the "positive form" for both the source and target of comparison, based on Greenberg's (2015, 2018) "gradability-based" semantics of *even*.

Finally, following Singh 2011, we proposed that where (20a) and (20b) compete, the EA comparative in both creates an updated context in which its presupposition is satisfied locally at the level of the subclause *more gigantic than Mothra*. This local context, in which the "positive form" presupposition is satisfied, causes the preference of (20b) with *even*, which scopes over and triggers the same presupposition as the EA comparative, due to a Maximize Presupposition!-type principle. The disfavored (20a) with the null form, which does not presuppose the "positive form," will consequently be degraded.

To summarize how our localized Maximize Presupposition!-like effect differs from traditional Maximize Presupposition!: (i) This effect operates locally, on the level of subclauses (Singh 2011) (or lexical items (Percus 2006)) and not globally. In sentences such as (18) and EA comparatives, this may lead to a situation where a sentence, which has two triggers for the same presupposition, is favored over a sentence with one trigger.⁵ (ii) This localized Maximize Presupposition-like effect is one where the disfavored competing expression is a null form, as in Amsili & Beyssade 2006.

⁵Thanks to Alexandre Cremers and Benjamin Spector, in attendance at CSSP 2017, for pointing out this issue.

4.5 Summarizing our Proposal and How it Solves the Two Puzzles

According to our proposal, the first puzzle (why are EA comparatives usually considered to be degraded?) is actually answered by our explanation to the second one (why does *even* improve the felicity of EA comparatives?). We derive the infelicity of (20a) from the fact that (20b) is preferred due to a local Maximize Presupposition!like effect.

Specifically, we proposed that, in fact, it is the absence of *even* (in (20a)) that causes the questionable felicity of the EA comparative due to the integration of three facts: (i) that with such comparatives, a presupposition of the "positive form" is taken to be satisfied; (ii) that there is a competing variant of such comparatives with *even*, which independently triggers a presupposition of the "positive form"; and (iii) that in such cases the presuppositionally stronger variant is favored, while the one with a weaker or without a presupposition is perceived as degraded, due to an effect similar to localized Maximize Presupposition!.

5 Supporting Evidence for Our Proposal

5.1 Cross-Linguistic Data on Cognates of English still

Our proposal that *even* makes EA comparatives felicitous by triggering the "positive form" presupposition predicts that the same effect should hold with other particles that trigger a similar presupposition.

A few cognates of *still* were observed in the literature to trigger a similar presupposition as part of their semantics or in comparative structures.⁶ These particles are, for example, French *encore* (as in (21); Hansen 2007), German *noch* (as in (22); Umbach 2009) and Hebrew *od* (as in (23); Greenberg 2012):

(21) Luc est **encore** plus beau qu'Adrien. (Hansen 2007:(114)) 'Luc is still better looking than Adrien.'

⁶Thanks to an anonymous EISS reviewer for pointing us in this direction.

Presupposes: Adrien is good-looking.

- Berta ist noch größer als Adam. (Umbach 2009:(4))
 'Berta is still taller than Adam.'
 Presupposes: Adam is tall.
- (23) Rina od yoter gvoha mi-Sara. (Greenberg 2012:fn. 6)
 'Rina is still taller than Sarah.'
 Presupposes: Sarah is tall.

Without going into the similarities and differences between these particles and *even*, the prediction of our proposal is that these particles, which trigger a similar presupposition to that triggered by *even*, would also greatly improve the felicity of EA comparatives. This prediction seems to be borne out. These particles, similarly to *even*, indeed appear to significantly improve the felicity of EA comparatives: (24b) with *encore*, (25b) with *noch* and (26b) with *od* seem to be more felicitous, compared to the degraded (24a), (25a) and (26a) without them:

(24) a. ??Rencontrer la France est plus énorme qu'affronter l'Angleterre.

'To encounter France is more enormous than to face England.'

- Rencontrer la France est encore plus énorme qu'affronter l'Angleterre.⁷
 'To encounter France is still more enormous than to face England.'
- (25) a. ??Seine Gelassenheit ist gigantischer als seine Technik. 'His serenity is more gigantic than his technique.'
 - b. Seine Gelassenheit ist **noch** gigantischer als seine Technik.⁸

'His serenity is still more gigantic than his technique.'

⁷https://tinyurl.com/yy5vw66k

⁸https://www.weltwoche.ch/ (registration required); then ausgaben/ 2006-37/artikel/artikel-2006-37-gross-gelassen.html

- (26) a. ??ha-Sulxan ha-kaxol yoter anak me ha-Sulxan ha-afor.'The blue table is huger/more huge than the grey table.'
 - b. ha-Sulxan ha-kaxol od yoter anak me ha-Sulxan ha-afor.
 'The blue table is still huger/more huge than the grey table.'

Preliminarily, this observation seems to strengthen the notion that there is a connection between the improved felicity of EA comparatives in the presence of *even* (and these other particles) and the "positive form" presupposition, which they and *even* share.

5.2 The Difference between Extreme and Lower-Closed Adjective Scales

As noted in section 4.1 above, Rett (2008) suggested that extreme adjectives have in fact lower-closed scale structure, based on an apparently similar presuppositional pattern, as illustrated by (27) (extreme adjective; repeated from (1b)) and (28) (lower-closed adjective; repeated from (11)):

- (27) ??Godzilla is more gigantic than Mothra. *Presupposes:* Mothra is gigantic.
- (28) This rag is wetter than the chamois. *Presupposes:* The chamois is wet.

However, note that the comparative in (27), with the extreme adjective, is, as already noted above, infelicitous (without *even*), while in (28) with a lower-closed scalar adjective, it is perfectly felicitous as it is. If, then, extreme adjectives are, in fact, associated with lowerclosed scales, this felicity difference would be problematic for our analysis. This because lower-closed scale structure, which has been argued to correlate to a minimum standard (cf. Kennedy & McNally 2005) can be seen as logically leading to the "positive form" presupposition, in the following way: if the standard is at the scale's minimal point, then even the smallest amount of wetness is considered to be [pos] wet, the same way the least amount of giganticness, is deemed to be [pos] gigantic. If both lower-closed and extreme adjectives indeed trigger the "positive form" presupposition, then one would expect (28), with the lower-closed-comparative, to violate the localized Maximize Presupposition as well, but as observed above, this is not the case.

It turns out that the explanation for this felicity difference ends up supporting our theory by highlighting an important difference between extreme adjectives and their non-extreme counterparts. While Rett (2008) claims that lower-closed scalar adjectives in the comparative presuppose the "positive form" regarding the source of comparison, as suggested in (28) above (cf. Demonte 2011), Kennedy (2007) argues that while there is such a presupposition with respect to the target of comparison, it is only a strong implication with respect to the source within the comparison. In sentences such as (28), for example, Kennedy claims that it is presupposed that the rag (target) is wet, but that it is only strongly implied that the chamois (source) is wet. This is, Kennedy proposes, because the chamois could, in principle, have "zero wetness" (i.e., be completely dry) while still being considered as having a degree on the wetness scale (see Kennedy 2007:fn. 23).

We adopt Kennedy's view and propose that the fact that the "positive form" is not presupposed, but only strongly implied for such adjectives in the comparative is why the conditions are not in place for a Maximize Presupposition!-like mechanism to take effect. Crucially, this is in opposition to the case with extreme adjectives. With the latter, as discussed in section 4.1 above, due to their inherent characterization as involving degrees which are just above the salient subscale *C*, it is necessarily presupposed that the source of the comparison has a degree which is at least as high as the membership standard.

Specifically, contrary to the case with lower-closed adjectives, with extreme ones it appears to be impossible to have a zero degree on the extreme (*gigantic*) subscale, because its lower bound is the largest degree on the non-extreme (*big*) subscale, as explained above. All

degrees above *C* (the salient portion) have a degree of giganticness which is above zero. Any degree below *C* is not on the giganticness scale anymore but on the bigness scale.

Thus, our proposal is supported by the difference between extreme adjectives and lower-closed scalar adjectives in that only in a context where the "positive form" is presupposed (as in EA comparatives) will the competing utterance with *even* be favored over the utterance without it, which will be degraded. In a case where the "positive form" is only strongly implied (as in lower-closed comparatives), *even* will not be required, because the conditions for the Maximize Presupposition!-like effect will not have been met.

6 Summary and Open Questions

In this paper we addressed two previously discussed puzzles concerning extreme adjectives in comparative constructions (EA comparatives). The first puzzle concerned the degraded status of such constructions, and the second concerned their improved felicity in the presence of *even*.

We reviewed Morzycki's (2012) intuitions and semantics of extreme adjectives, and how these are meant to address the two puzzles, but pointed out a few challenges for these suggestions. Instead, we developed a proposal, where we followed Morzycki in assuming that the interaction between the semantics of comparatives and his proposed scale structure of extreme adjectives, causes EA comparatives to trigger a "positive form" presupposition, where the source of comparison (as well as the target) has a degree which is at least as high as the standard for that extreme adjective.

As part of our proposal, we relied on the independently motivated "gradability-based" analysis of *even* (Greenberg 2015, 2018). This, in turn, left us with two components in the structure that trigger the "positive form" presupposition with respect to the source: the EA comparative itself, based on Morzycki's analysis, and *even*, based on the "gradability-based" semantics.

We then proposed that a mechanism akin to local Maximize Pre-

supposition! causes the preference of an utterance in which *even* combines with the EA comparative, since then the "positive form" presupposition triggered by *even* is independently assumed to be triggered and locally satisfied by the EA comparative. As a result, the variant with *even* is favored over the competing variant without *even*, which is thus judged to be degraded.

We provided two supporting arguments for our proposal. The first one is a prediction of our theory, which appears to be borne out. Particles in French, German and Hebrew, which have been claimed to have a similar "positive form" presupposition, also seem to have an improving effect on EA comparatives. A second prediction of our analysis, which appears to be borne out, is that only an environment in which the "positive form" is presupposed (as with EA comparatives) would cause the Maximize Presupposition!-like effect to come into play, as opposed to where the same inference is only strongly implied (as in with lower-closed scalar adjectives, which are perfectly fine in the comparative).

Hopefully, our analysis can help to shed light on the nature of extreme adjectives, their associated scale structure and their membership standard, as well as to contribute to research on other types of adjectives, Maximize Presupposition!, and the semantics of *even*.

The proposal above leaves a number of questions open and raises possible directions for future research. Beginning with the open questions, first, the status of EA comparatives without *even* is often judged to be degraded but not totally infelicitous like the status of, for example, (16a) (#*A sun is shining*), which is supposedly blocked by Maximize Presupposition!. We would like to examine whether this difference can be attributed to a competition between an overt and a null form that we proposed occurs with EA comparatives and *even*,⁹ as

⁹It bears mentioning that the proposal by Amsili & Beyssade (2006) that we used to support our proposal, whereby *too* competes with a null form within Maximize Presupposition!, seems to face counterexamples. Consider (i):

⁽i) a. Jean est malade, mais je ne pense pas que Marie le soit aussi.

opposed to a competition between two overt forms, as in the classic examples of Maximize Presupposition!.

In addition, we found various attested examples where EA comparatives appear to be felicitous, also without the presence of *even*. A preliminary online stock-taking of such cases appears to indicate that there are at least two groups of such examples, one involving NPIs as in (29), and another where no *than*-phrase is explicitly used (or when the comparative is used attributively), as in (30).

- (29) a. America astonished him with social paradoxes far more gigantic than **anything** he had observed in Europe.¹⁰
 - b. A large tree far more ginormous than Bresa could **ever** imagine reached towards the sky in the middle of this odd valley.¹¹
 - c. (...) the laundry pile is more ginormous than **ever**.¹²
 - With cakes going more ginormous and extreme than ever before, ever wonder what the world's largest cake looks like?¹³
- (30) a. People who wear crazy socks are more brilliant, creative

'John is sick, but I don't think Mary is too.'

b. Jean est malade, mais je ne pense que Marie le soit. 'John is sick, but I don't think Mary is.'

Sentence (ib), with *too*, is felicitous and the presupposition that someone else in the context, namely John, is sick as well, holds. Based on Amsili & Beyssade's proposal, *too* is predicted to be obligatory in such a scenario, because it triggers a stronger presupposition that holds in the context compared to the null form. However, (ia) without the presuppositionally stronger *too*, is perfectly felicitous. It merits further research to check the viability of an overt vs. covert form competition within Maximize Presupposition!. Thanks to an anonymous EISS reviewer for pointing out this issue and this example.

¹⁰Google books search result; quote from: Woodcock, George. 2007. *Dawn and the darkest hour: A study of Aldous Huxley*. Vol. 350. Black Rose Books Ltd.

¹¹https://tinyurl.com/yynb5ut2
¹²https://tinyurl.com/yy42erc9
¹³https://tinyurl.com/y6teygt6

and successful.14

- The forest seems to decrease in height in these calcareous rocks, especially the planes, which are more colossal in Indiana.¹⁵
- c. When I hear "advanced" preferences, I think of more miniscule nuances such as espresso or French roast.¹⁶
- d. The palate may want more sumptuous food or, at times, a dish that requires more elaborate preparation.¹⁷

More research is needed to check whether our analysis can account for these cases in a precise manner.

Beyond the examples in (29) and (30), there appear to be (at least borderline) felicitous sentences without *even*, in which the "positive form" clearly holds of the source of comparison and which would be felicitous with *even*. These examples potentially pose a greater problem for our proposal, which would predict them to be degraded. Consider (31) and (32).

- (31) A: John is tall. He is 1.80m tall.B: And what about Bill?A: He is (even) taller (than John). He is 1.86m tall.
- (32) Godzilla is (even) as gigantic as Mothra.

In the exchange in (31), the "positive form" holds of John (the source) in the context, but *even* appears to be optional and not obligatory. Similarly, the equative in (32), which like the comparative, also presupposes 'Mothra is gigantic', does not require the presence of *even*.

¹⁴Thanks to an anonymous EISS reviewer, who pointed out this example: https://tinyurl.com/y34tyvzf

¹⁵Google books search result; quote from: Wied, Maximilian. 1843. *Travels in the interior of North America*. Ackermann and Company.

¹⁶Google books search result; quote from: Potts, Kevin et al. 2007. *Textpattern* solutions: PHP-based content management made easy. Apress.

¹⁷Google books search result; quote from: Symes, Carol. 2008. *Abelard and Heloise: The letters and other writings*. Hackett Publishing Company.

While it is not entirely clear that (31) is not improved with *even*, or how exactly the equative interacts with extreme adjectives and *even* in (32), these two and other examples like them, definitely merit further examination.

One interesting direction for further research concerns the interaction of extreme adjectives with *less* ... *than* comparatives,¹⁸ illustrated in (33).

(33) Godzilla is ??(even) less gigantic than Mothra.

Surprisingly, it appears that the effect observed with *even* and *more* ... *than* EA comparatives is reversed with *less* ... *than*. Here, the variant without *even* appears to be acceptable, while the version with *even* appears to be odd. It would be interesting to check whether the "gradability-based" semantics of *even* and the above proposal as a whole, can be used to explain this effect.

A final direction for further research involving a different solution to the puzzles addressed by this proposal, involves Morzycki's idea regarding the pragmatic clash caused by comparing degrees of extreme adjectives (described in section 2.2.1). Recall that Morzycki proposes that extreme adjectives inherently involve degrees which are non-salient, and that this attribute clashes with the act of comparing, which automatically "brings degrees into salience," thus causing the observed infelicity. This intuitively appealing explanation may be perhaps useful if defined in terms of informativity or relevance towards resolution of a QUD. Perhaps it is possible to claim, roughly, that the distinction between two extreme degrees is non-informative or non-relevant for the QUD (following Portner & Rubinstein 2016). Preliminarily, the improved felicity with even may be explained by the fact that the latter's "gradability-based" semantics defines the scale associated with it as a contextually relevant scale. This way, even indicates that in its presence, a comparison of extreme degrees

¹⁸Thanks to an anonymous IATL 2017 reviewer for suggesting this line of inquiry.

is informative, thus countering the initial inherent non-informativity of such a comparison. This idea requires further research and elucidation.

Acknowledgments We would like to thank Alexandre Cremers, Regine Eckardt, Markus Egg, Carlo Geraci, Jeremy Kuhn, Galit Sassoon, Benjamin Spector, Malte Zimmermann and anonymous EISS and IATL reviewers for their constructive comments.

Research on this paper was supported by ISF grant 1655/16, given to Yael Greenberg.

References

- Amsili, P., & Beyssade, C. (2006). Compulsory presupposition in discourse. In C. Sidner, J. Harpur, A. Benz, & P. Kühnlein (Eds.) *Proceedings of the* 2nd Workshop on Constraints in Discourse, (pp. 5–10). Maynooth: National University of Ireland.
- Chemla, E. (2008). An epistemic step for anti-presuppositions. *Journal of Semantics*, 25(2), 141–173.
- Chierchia, G. (2013). *Logic in Grammar: Polarity, Free Choice, and Intervention*. Oxford University Press.
- Demonte, V. (2011). Adjectives. In K. von Heusinger, C. Maienborn, & P. Portner (Eds.) *Semantics: An International Handbook of Natural Language Meaning*, vol. 2, (pp. 1314–1340). Berlin: De Gruyter Mouton.
- Giannakidou, A. (2007). The landscape of *even*. *Natural Language and Linguistic Theory*, 25(1), 39–81.
- Greenberg, Y. (2012). Event-based additivity in English and Modern Hebrew. In P. C. Hofherr, & B. Laca (Eds.) *Verbal Plurality and Distributivity*, Linguistische Arbeiten, (pp. 127–159). Berlin: De Gruyter.
- Greenberg, Y. (2015). *Even*, comparative likelihood and gradability. In
 T. Brochhagen, F. Roelofsen, & N. Theiler (Eds.) *Proceedings of Amsterdam Colloquium*, vol. 20, (pp. 147–156).
- Greenberg, Y. (2018). A revised, gradability-based semantics for *even*. *Nat*-*ural Language Semantics*, 26(1), 51–83.
- Guerzoni, E. (2003). Why Even Ask? On the Pragmatics of Questions and the Semantics of Answers. Ph.D. thesis, MIT.
- Hansen, M.-B. M. (2007). Particles at the Semantics/Pragmatics Interface. Lei-

den, The Netherlands: BRILL.

- Heim, I. (1991). Artikel und Definitheit. In A. von Stechow, & D. Wunderlich (Eds.) *Semantik: Ein internationales Handbuch der zeitgenössischen Forschung*, (pp. 487–535). Berlin: De Gruyter.
- Horn, L. (1969). A presuppositional theory of 'only' and 'even'. *Chicago Linguistics Society*, *5*, 97–108.
- Karttunen, L., & Peters, S. (1979). Conventional implicature. In D. A. Dinneen, & C.-K. Oh (Eds.) *Presupposition*, Syntax and Semantics 11, (pp. 1–56). New York: Academic Press.
- Kennedy, C. (2004). Comparatives, semantics of. In A. Barber, & R. Stainton (Eds.) Concise Encyclopedia of Philosophy of Language and Linguistics, (pp. 68–71). Oxford: Elsevier.
- Kennedy, C. (2007). Vagueness and grammar: The semantics of relative and absolute gradable adjectives. *Linguistics and Philosophy*, 30(1), 1–45.
- Kennedy, C., & McNally, L. (2005). Scale structure, degree modification, and the semantics of gradable predicates. *Language*, *81*(2), 345–381.
- Morzycki, M. (2012). Adjectival extremeness: Degree modification and contextually restricted scales. *Natural Language and Linguistic Theory*, 30(2), 567–609.
- Paradis, C. (1997). Degree modifiers of adjectives in spoken British English. In *Lund Studies in English*, vol. 92. Lund University Press.
- Paradis, C. (2001). Adjectives and boundedness. *Cognitive Linguistics*, 12(1), 47–65.
- Percus, O. (2006). Anti-presuppositions. In A. Ueyama (Ed.) *Theoretical and Empirical Studies of Reference and Anaphora: Toward the Establishment of Generative Grammar as an Empirical Science*, (pp. 52–73). Japan Society for the Promotion of Science.
- Portner, P., & Rubinstein, A. (2016). Extreme and non-extreme deontic modals. In N. Charlow, & M. Chrisman (Eds.) *Deontic Modals*, (pp. 256–282). Oxford University Press.
- Rett, J. (2008). *Degree Modification in Natural Language*. Ph.D. thesis, Rutgers University.
- Rooth, M. (1985). *Association with Focus*. Ph.D. thesis, University of Massachusetts, Amherst.
- Rooth, M. (1992). A theory of focus interpretation. *Natural Language Semantics*, 1(1), 75–116.

- Rullmann, H. (1997). *Even*, polarity, and scope. In M. Gibson, G. Wiebe, & G. Libben (Eds.) *Papers in Experimental and Theoretical Linguistics*, vol. 4, (pp. 40–64). University of Alberta.
- Rullmann, H. (2007). What does *even* even mean. Handout, University of Calgary, 2007/12/07.
- Sauerland, U. (2008). Implicated presuppositions. In A. Steube (Ed.) *The Discourse Potential of Underspecified Structures*, vol. 8, (pp. 581–600). Leipzig: De Gruyter.
- Schlenker, P. (2012). Maximize Presupposition and Gricean reasoning. *Natural Language Semantics*, 20, 391–429.
- Singh, R. (2009). Maximize Presupposition! and informationally encapsulated implicatures. In A. Riester, & T. Solstad (Eds.) *Proceedings of Sinn und Bedeutung*, vol. 13, (pp. 513–526). Stuttgart: Online Publikationsverbund der Universität Stuttgart.
- Singh, R. (2011). Maximize Presupposition! and local contexts. *Natural Language Semantics*, 19(2), 149–168.
- Umbach, C. (2009). Another additive particle under stress: German additive *noch*. In L. Kálmán (Ed.) *Proceedings of the Tenth Symposium on Logic and Language*, vol. 10, (pp. 149–156). Budapest: Eötvös Loránd University.
- Wilkinson, K. (1996). The scope of *even*. *Natural Language Semantics*, 4(3), 193–215.

(Non-)Exhaustivity in French c'est-Clefts

Emilie Destruel • Joseph P. De Veaugh-Geiss

Abstract This paper presents two experimental studies that used an incremental information-retrieval paradigm to compare the exhaustivity in *c'est*-clefts to exhaustivity inferences in other constructions in French, as well as to exhaustivity in comparable constructions in German. Results suggest that exhaustivity in *c'est*-clefts is weak and in some cases shows a divergent pattern from exhaustivity in definite pseudoclefts (that is, identity statements with a definite description), contra predictions of Percus (1997) and Büring & Križ (2013) and differing from an identical study on German (De Veaugh-Geiss et al. 2017, 2018). We seek a unified account of cleft exhaustivity, and propose that the broader discourse semantics of *c'est*-clefts accounts for their weak exhaustivity effect.

Keywords c'est-clefts \cdot es-clefts \cdot exhaustivity \cdot experimental study \cdot German \cdot French

E. Destruel, University of Iowa, https://www.emiliedestruel.org/ J. P. De Veaugh-Geiss, Universität Potsdam, https://www.ling.uni-potsdam.de/~deveaugh-geiss/

In Christopher Pinon (Ed.), Empirical Issues in Syntax and Semantics 12, 91–120. Paris: CSSP. http://www.cssp.cnrs.fr/eiss12/ © 2019 Emilie Destruel & Joseph P. De Veaugh-Geiss

1 Introduction

Similar to other cross-linguistic focus-background constructions (e.g., the English *it*-cleft and the German *es*-cleft), the French *c'est*-cleft in (1) is claimed to have three standard components. It conveys a prejacent proposition that amounts to the corresponding canonical form (1a). It carries an existence presupposition (1b). Finally, it gives rise to an exhaustive inference whereby the predicate holds for no individual other than the one denoted by the cleft pivot (1c).

(1) C'est Marc qui a préparé un cocktail.'It is Marc who prepared a cocktail.'

- a. Marc prepared a cocktail.
- b. Someone prepared a cocktail.

(prejacent) (existence)

c. Nobody other than M. prepared a cocktail. (*exhaustive*)

It is commonly accepted that the prejacent is part of the asserted content, while the existential is presupposed. The source of exhaustivity is, however, disputed, and this debate is still very much alive, with broadly speaking—two main theoretical positions offered to explain exhaustivity in clefts, either in semantic or in pragmatic terms. Experimental work, however, suggests that exhaustivity in clefts is generally not strongly conveyed (see, e.g., Destruel 2013 for French; De Veaugh-Geiss et al. 2018 for German), and for French specifically, recent studies found *c'est*-clefts to differ from English clefts in exhibiting weaker exhaustivity effects (Destruel & De Veaugh-Geiss 2018).

This paper seeks to address the following: how does exhaustivity in French clefts compare to (i) exhaustivity in clefts in other languages, in particular, German, and to (ii) other exhaustive inferences, in particular, that conveyed in definite pseudoclefts (that is, identity statements with a definite description)? To this end, we ran two experiments identical to recent studies on German (De Veaugh-Geiss et al. 2018) employing an incremental information-retrieval paradigm. The task involved a mouse-driven falsification/verification task in which participants had to make a truth-value judgment for target sentences conveying exhaustivity, including exclusives, narrow (prosodic) subject-focus, definite pseudoclefts, and clefts. The main contribution of the paper is that it provides additional evidence to the rather small empirical literature on exhaustivity in French, and expands on a prior experimental design to compare exhaustivity cross-linguistically.

2 Background

Exhaustivity is not specific to clefts, but is also conveyed by other sentence forms, most notably exclusives (2a), in situ narrow (prosodic)

subject focus (2b), and definite pseudoclefts (2c).

| (2) | a. | Seul Marc a préparé un cocktail. | | |
|-----|----|----------------------------------|-------------|--|
| | | 'Only Marc prepared a cocktail.' | (exclusive) | |
| | b. | MARC a préparé un cocktail. | | |
| | | 'MARC prepared a cocktail.' | (focus) | |

c. La personne qui a préparé un cocktail est Marc. 'The person who prepared a cocktail is Marc.'

Exhaustivity is not derived the exact same way in each of these sentence forms. There is a general consensus in the literature on the fact that, with exclusives, exhaustivity is part of the asserted content of the sentence. It is also uncontroversial that exhaustivity is typically obtained via pragmatic enrichment in sentences with (prosodic) narrow focus. Definite pseudoclefts, by contrast, have been argued to share the underlying syntactic structure (Percus 1997) and semantic contribution of focus-background *it*-clefts (Percus 1997, Büring & Križ 2013, Hedberg 2013), and under these analyses exhaustivity is a hard-coded but not-at-issue maximality or homogeneity presupposition. The question concerning how to theoretically model exhaustivity in clefts nevertheless remains hotly debated. We now offer a brief overview of the debate, largely based on English *it*-clefts.

2.1 Theoretical Accounts and Empirical Findings

Past theoretical accounts of cleft exhaustivity generally split along a semantic/pragmatic divide. In a nutshell, this comes down to the question of whether exhaustivity is conventionally coded as part of the cleft's meaning (semantic accounts) or not (pragmatic accounts). Although mainly developed with a view to English, it is important to note that the analyses reviewed hereafter can arguably be applied to other languages in which cleft exhaustivity has been acknowledged as well (e.g., French, German, etc.)

While early semantic approaches analyzed clefts on par with exclusives (Atlas & Levinson 1981, É. Kiss 1998), later analyses are

⁽def. pseudocleft)

less strict, claiming that exhaustivity in clefts is in some way presupposed. Within this line of argument, scholars have exploited the similarity between clefts and definite descriptions; see, for example, Percus (1997) and Hedberg (2013) for an account specified in terms of a maximality presupposition, and Büring & Križ (2013) for an analysis in terms of a homogeneity presupposition. Others, such as Velleman et al. (2012), have taken clefts to be *inquiry-terminating* devices, that is, devices that give a final and complete answer to a question. Under this view, the exhaustive component of clefts has a different discourse-semantic status than that in exclusives: it is not at-issue. In other words, exhaustivity is not the main point of what is conveyed by a cleft, rather the prejacent proposition is—and vice versa for exclusives.

Crucially, all semantic accounts contend that the inference is derived directly from the linguistic form of the utterance, making the following predictions: cleft exhaustivity is **systematic** and **robust**, and therefore not (easily) cancellable. It is predicted to arise across experimental manipulations and across speakers. Moreover, according to certain semantic approaches (Percus 1997, Büring & Križ 2013), it is expected to arise in a parallel fashion to definites.

By contrast, the pragmatic view, largely advocated by Horn (1981, 2014), takes exhaustivity to simply be added to the meaning of the sentence as a (generalized) conversational implicature, and this based on the observation that exhaustive effects do not seem obligatory with clefts (see Horn 1981; Horn 2014 for English; Destruel 2013 for French; De Veaugh-Geiss et al. 2015 for German). For instance, De Veaugh-Geiss et al. (2015) present a pragmatic analysis of exhaustivity in clefts in which exhaustivity is a focus-triggered scalar implicature. They argue that the differences between clefts and canonicals in terms of cancellation possibilities are not due to a particularly strong exhaustivity effect in clefts *per se*—unlike the manner implicature account in Horn 1981—but rather due to a weak exhaustivity effect in plain focus constructions given potential projection ambiguities (argued for even narrow subject focus) and thus suboptimal

environment for further pragmatic enrichment (see, however, Križ 2017: 5–6 for arguments against such an approach).

All in all, pragmatic accounts make straightforward empirical predictions contrasting with those made by semantic accounts: cleft exhaustivity is subject to **defeasibility** (i.e., it is not robust) and **variation** across contexts (i.e., it lacks systematicity). However, assuming that the derivation of exhaustivity follows **universal** pragmatic principles, we should expect minimal cross-linguistic variation.

Positioned outside of this semantic-pragmatic divide, the dynamic account of Pollard & Yasavul (2014), which following De Veaugh-Geiss et al. (2018) we draw from in our analysis, does not take exhaustivity to be coded in the cleft, but rather to be the result of the interaction of the existence presupposition of clefts (1b) with the meaning of *wh*-questions (Hamblin 1973). In this account, clefts specify an antecedent discourse referent, and this in two ways. In the non-exhaustive case, clefts pick up some (non-maximal) discourse referent to designate further. This can be illustrated in the case of, for example, correction, when revising misinformation about a referent in the discourse, as in (3).¹

(3) A: Did you hear, Bob got an NSF grant!
B: Well, actually, it was Peter. And Mike got one, too! (Adapted from ex. (5) in Pollard & Yasavul 2014)

When clefts answer *wh*-questions, however, an exhaustivity implication arises: the question introduces a maximal discourse referent, and the cleft existential has this discourse referent as its antecedent.

Recent years have seen an increase in experimental work testing

- (i) A: Did you hear, Bob got an NSF grant!
 - B: Well, actually, only Peter got an NSF grant. #And Mike got one, too!

¹One could claim that the acceptability of the second clause in B's response is an example of domain widening; however, in the same discourse with the exclusive *only* instead of the cleft, the continuation becomes infelicitous. Arguably, domain widening should apply in such cases as well.

predictions from various sides of the debate. One influential study by Onea & Beaver (2009) involved a violation task on Hungarian preverbal focus, later replicated for clefts (with comparable results) in English, German, French, and Greek (see Xue & Onea 2011, Destruel 2012, Destruel et al. 2015, Pavlou 2015). These studies found that in cases when exhaustivity was violated, participants chose a weak *Yes, but* continuation for preverbal focus and clefts, suggesting that exhaustivity in clefts was less robust than predicted under a semantic account. Similar violation diagnostics were applied in acceptability and truth-value judgment studies (e.g., De Veaugh-Geiss et al. 2018). Results again showed that, despite the explicit violation of exhaustivity, clefts were judged as relatively acceptable, and thus exhaustivity was argued not to be coded in the cleft structure itself.

To sum up: although all accounts embody the notion that exhaustivity is present with clefts, semantic approaches predict a systematic and robust link between clefts and the exhaustive inference (especially in unembedded contexts where presuppositions normally cannot be cancelled or suspended), with similar behavior for definites expected under certain proposals. Pragmatic approaches generally predict the opposite. Recent experimental work has mostly challenged strict semantic analyses based on the finding that the strength of the exhaustive inference is rather weak and variable.

2.2 The Case of French c'est-Clefts

There is some support for suggesting that French *c'est*-clefts are semantically similar to the focus-background clefts commonly discussed in the literature. Indeed, several French scholars have noted that *c'est*-clefts come with an existence presupposition and convey exhaustivity effects (Lambrecht 1994, Katz 2000, De Cat 2007). However, while exhaustivity in *c'est*-clefts is acknowledged within the literature on French proper (Lambrecht 1994, 2001, Katz 1997, De Cat 2007), few researchers have directly addressed the issue of how it is derived.

Drawing upon analyses in the cross-linguistic literature, Clech-

Darbon et al. (1999) are among these few arguing for a truth-conditional account (citing É. Kiss 1998) under which exhaustivity in clefts is equated to that in exclusives. Most recently, Destruel (2013) follows Horn (1981) in arguing for an alternative pragmatic view, namely, an implicature-based account. Although empirical work on French is quite scarce, Destruel (2013) and Destruel et al. (2015) suggest that *c'est*-clefts are only weakly exhaustive—and to a much lesser extent than exclusives.

Indeed, nothing precludes extending past accounts to French, and yet, there are also some subtle differences that set this language apart. First, *c'est*-clefts are the most common, and thus unmarked, strategy to signal subject focus (Lambrecht 1994, Carter-Thomas 2009, Féry 2013), and thus they have a high frequency in French (see, e.g., Dufter 2009 for a cross-linguistic perspective with corpus data on Romance languages vs. German). This has been argued to be primarily due to constraints on French prosody: whereas in Germanic languages, for example, English or German, prosodic prominence can be shifted to match the location of the focus constituent, French is more rigid, placing prosodic stress only at the right edge of an intonation phrase. The *c'est*-cleft, despite adding syntactic complexity, circumvents this prosodic restriction by creating an extra intonation boundary that can align with the focused constituent (Hamlaoui 2009).

Second and most importantly, French *c'est*-clefts have a broad discourse function: in addition to signaling a narrow focus, they can be used in all-focus contexts, for example, contexts in which the focus falls on the entire sentence. This is evident in examples such as (4) from Clech-Darbon et al. (1999), in which the answer to the question for the cleft of the form *It is X who P* is not congruent with a question derived from the cleft relative, that is, *who P*?—or a subquestion of this question, for example, *which x P*?—but rather, the much broader question *What happened*?²

²Recent corpus studies have provided further evidence of the occurrence of this type of *c'est*-cleft (Karssenberg & Lahousse 2015).

- (4) Q: Qu' est-ce qui s' est passé? what is-it that REFL.3SG is happened 'What happened?' (Lit.) 'What is it that happened?'
 - A: C' est le petit qui est tombé dans l' escalier. It is the small-one who is fallen in the stairs 'The little one fell down the stairs.' (Lit.) 'It is the little one that fell down the stairs.'

In contrast, and in line with De Veaugh-Geiss et al. (2015), projection from the cleft pivot to constituents outside of the pivot appears to not be possible in German or English—and thus in these languages only narrow focus is possible. Thus, the question corresponding to a cleft matches the focus-background structure of the cleft-structure (i.e., it is directly derivable from the cleft relative clause); or for complex cleft pivots, a subquestion of the question delimited by the cleft relative, disambiguated by focus-marking (Velleman et al. 2012, Križ 2017). This leads to a (semi-)strict relationship in these languages between the cleft and the question it can answer (Abrusán 2016).

Empirically, a recent study by Tieu & Križ (2017) on the L1 acquisition of exhaustivity hints at differences between English and French clefts. Existing data on the acquisition of English *it*-clefts suggests that children start out by interpreting clefts non-exhaustively, and have partly acquired exhaustivity around the age of 4–5 years old (Heizmann 2007, 2012). In Tieu & Križ's (2017) truth-value judgment task, children looked at pictures containing three familiar objects (created in an exhaustive and a non-exhaustive condition) while a puppet described them in a video using a cleft sentence (among others). Children were asked to judge whether what the sentence uttered by the puppet accurately described the picture or not. Although, like in English, French-speaking children start out by interpreting clefts non-exhaustively, they were found to continue interpreting clefts non-exhaustively at 6 years old (i.e., comparatively later than English-speaking children in Heizmann's studies).

(Non-)Exhaustivity in French c'est-Clefts

To date, though, virtually no studies have directly compared the exhaustive effects in French clefts versus clefts in other languages. One exception is Destruel & De Veaugh-Geiss (2018), who addressed this deficit by testing the differences in interpretation and processing costs of exhaustivity in French vs. English clefts. Results overall indicate that exhaustivity in *c'est*-clefts is weaker than in *it*-clefts, and that, while English clefts elicited higher processing costs compared to exclusives/canonicals when exhaustivity was violated, French clefts did not. Taken together, these findings provide a more nuanced cross-linguistic picture of cleft exhaustivity, which we aim to complement with the experiments reported in the next section.

3 The Experiments

3.1 Research Questions

The specific questions addressed in this study are:

- (i) How does the systematicity and robustness of exhaustivity in French clefts compare to exhaustivity effects reported in an identical study on German clefts?
- (ii) Are there parallels between exhaustivity in French *c'est*-clefts and other sentence types, in particular definite pseudoclefts, as reported for German *es*-clefts (De Veaugh-Geiss et al. 2018)?

3.2 Methods and Design

To test the systematicity and robustness of exhaustivity in French clefts, and crucially to be able to compare the results directly to those currently present in the literature for other languages, we adapted for French a design developed in a recent study on German, namely a mouse-guided incremental information retrieval paradigm with a verification (Experiment I) and falsification (Experiment II) task (De Veaugh-Geiss et al. 2017, 2018).³

³Although the methods and design are identical, we refer the reader to those sources for specifics regarding the German experiments. Also, note that several of the lexicalizations of the items differed; see appendix A and the supplementary materials with this paper for French, and the supplementary materials included

Participants For Experiment I we present the results for 32 French native speakers (20 female, 12 male; average age: 43.2; age range: 32–54), and for Experiment II 32 different French native speakers⁴ (18 female, 14 male; average age: 40.2; age range: 24–59), all students and staff at the University of Albi or the University of Pau, France.

Procedure and Design The experiments took part in a quiet room and were run on a computer using Python scripts (v.3.4.2 on GNU/Linux; v.3.3.5 on MS Windows) with the PyGame module (v.1.9.2ao, LGPL, Shinners 2011). Because the timeline and the experimental material for the joint experiments are identical, we will present them together and will emphasize the relevant differences when appropriate.

Before the experiments started, participants first read a set of instructions introducing them to four roommates: Charles, Pierre, Marc, and Jean. They were told that these four roommates, and these alone, were involved in activities together to be described in the experiment. At the beginning of each trial, participants saw four covered boxes on a computer screen, as in the left panel of figure 1, while hearing the target stimulus in their headphones. After the audio stimulus played in their headphones, participants were instructed to uncover as many boxes as necessary, one at a time by moving the mouse over it, in order to decide whether the sentence they heard was true or false. The right panel of figure 1 illustrates the uncovering of the critical second box (Box 2).

A 2000-ms pause was implemented between each box to discourage participants from uncovering boxes unnecessarily. Under each box was a picture of one of the four roommates and a written description of the activity he carried out. As soon as participants moved the cursor away from a box, the description disappeared but the im-

with De Veaugh-Geiss et al. 2018 for German.

⁴There were in fact 37 total participants in Experiment II, but 5 were removed from the final dataset for having less than 75% accuracy (i.e., choosing to continue despite the violation of exhaustivity) in the control condition for exclusives.



Figure 1 Onset screen (left), and uncovering of Box 2 (right)

age remained visible. Participants were allowed to move the cursor back to an uncovered box at any point during the trial if they wanted to see the description again. Finally, although participants were free to choose which box they uncovered next, it is important to note that the experiment was programmed with a pre-determined order.

Using an incremental information paradigm we are able to measure at which point participants had enough information to make a judgment about the target sentence they heard. Of the four boxes participants uncovered, there were two points of interest for both experiments: the Early Response measure (Box 2) and the Late Response measure (alternating between Box 3 and Box 4, depending on the experimental trial). The early response had three possible values; that is, whether a truth-value judgment was made (*vrai* 'true' or *faux* 'false'), or whether participants chose to continue to uncover the subsequent Box 3/4. The late response refers to the final evaluation given when considering all relevant information in cases when participants chose to continue at Box 2.

For the early response measure, we manipulated how Box 2 (i.e., the critical box in both experiments) related to EXHAUSTIVITY, and the primary difference between Experiment I (Verifier) and Experiment II (Falsifier) lies in this factor. In Experiment I, Box 2 always **verified** the prejacent or canonical meaning of the sentence, and hence the name "Verifier"; for example, Marc in fact claims he prepared a cocktail, for example, illustrated in (5).

(5) Target: C'est MARC qui a préparé un cocktail.'It is Marc who prepared a cocktail.'

- a. Box 1: Jean 'I served a drink.'
- b. Box 2: Marc 'I prepared a cocktail.'
- c. Box 3: Charles 'I tasted a whiskey.'
- d. Box 4: Pierre 'I prepared a cocktail.'

By comparison, in Experiment II, Box 2 always **falsified** the exhaustivity inference, and hence the name "Falsifier"; for instance, participants discover at the second box that someone other than Marc, for example, Jean, claims that he prepared a cocktail (not illustrated here for the sake of space; see the supplementary materials for details). It is important to note that Box 1 was always irrelevant for exhaustivity.

Materials Both experiments involve a 4x2 design: *Sentence* form of the auditory target stimuli (4 levels: EXCLUSIVE, FOCUS, DEFINITE PSEUDOCLEFT, CLEFT) and the late response controls, that is, *Exhaustivity* in Experiment I and *Canonical* in Experiment II (both 2 levels: +/–).

The first factor we will discuss, *Sentence* form, included CLEFTS, as in (1), the EXCLUSIVE control condition, as in (2a), canonical sentences with prosodic subject FOCUS, as in (2b), and DEFINITE PSEUDOCLEFTS (with definite descriptions of the form *la personne* 'the person'), as in (2c). As discussed, these four sentence forms have been claimed to associate with an exhaustive inference, amounting to '*nobody other than Marc prepared a cocktail*' (with the domain fixed to the four roommates Marc, Charles, Pierre, and Jean throughout the experiment).

The second factor, *Exhaustivity* (Experiment I) or *Canonical* (Experiment II), specifically involved the late response measures at Box 3/4, which served as additional controls that participants understood the logic of the experiments.

• In Experiment I, in which Box 2 verified the prejacent, at Box 3/4 in half the trials either (i) no one else is revealed to have prepared a cocktail (+EXH), thus satisfying exhaustivity, or (ii) someone other than Marc, for example, Pierre, is revealed to

have prepared a cocktail (-EXH), thus violating exhaustivity.

• In Experiment II, in which Box 2 falsified the exhaustivity inference, at Box 3/4 in half the trials either (i) Marc is revealed to have prepared a cocktail (+CAN), verifying the canonical meaning of the sentence, or (ii) Marc did something other than prepare a cocktail (-CAN), violating the canonical meaning.

An overview of the conditions can be found in the summary in table 1 (adapted for French from De Veaugh-Geiss et al. 2018).

| | Experiment I (Verifier) | Experiment II (Falsifier) | |
|------------------|--|----------------------------------|--|
| Box 1 | (irrelevant information) | | |
| | Jean: "I served a drink." | | |
| Box 2 | (canonical verified) | (exhaustivity falsified) | |
| (Early Response) | Marc: "I prepared a cocktail." | Pierre: "I prepared a cocktail." | |
| Box 3 / Box 4 | [+EXH] (exhaustivity verified) | [+CAN] (canonical verified) | |
| (Late Response) | Charles/Pierre: "I had a whiskey." | Marc: "I prepared a cocktail." | |
| | or | or | |
| | [-EXH] (exhaustivity falsified) | [-CAN] (canonical falsified) | |
| | Charles/Pierre: "I prepared a cocktail." | Marc: "I had a whiskey." | |

Audio: C'est Marc qui a préparé un cocktail. 'It is Marc who prepared a cocktail.'

 Table 1
 Example conditions of Experiment I (Verifier) & Experiment II (Falsifier).

There were 32 auditory target stimuli and 32 auditory filler items, the latter including sentences with universal quantifiers *tout le monde* 'everybody', expletive constructions beginning with *c'est clair que* 'it is clear that', plural conjunctions, and scalar constructions with *moins de trois personnes* 'fewer than three people'. The 64 total sentences, all with unique lexicalizations, were randomized during presentation.

3.3 Predictions

Predictions for the exhaustivity inferences in the sentence types tested are as follows. If exhaustivity is **semantic**—that is, conventionally coded, with EXCLUSIVES as the control condition—it will be robust and systematic, and thus:

- In Experiment I, in which the canonical meaning or prejacent has been verified at Box 2, a majority of participants will chose 'continue' to check that exhaustivity holds.
- In Experiment II, in which exhaustivity is falsified at Box 2, given the violation of the necessary entailments a majority of participants will make a 'false' judgment without uncovering Box 3/4.⁵
- However, cross-linguistic differences may emerge, since not all languages may encode exhaustivity in the same way.⁶

By contrast, should exhaustivity be **pragmatic**—that is, neither conventionally coded nor truth-functional, with FOCUS as the control condition—it will be defeasible and variable, and thus:

- In Experiment I, verifying the canonical meaning or prejacent may be sufficient to make a judgment, and participants can already make a 'true' judgment at Box 2 (although continuing to uncover Box 3/4 is also possible).
- In Experiment II, falsifying the exhaustivity inference could be insufficient for a final judgment, and participants may continue to uncover Box 3/4 in order to check the yet-unverified prejacent (although making a 'false' judgment is also possible).
- However, if (and only if), the exhaustive inference is derived by universal conversational principles, we predict minimal cross-linguistic variation.

3.4 Data Preparation and Analysis

For data preparation, responses at Box 2 were coded as 1 for judgment (Experiment I: 'true', Experiment II: 'false') and 0 for 'con-

⁵We assume that if exhaustivity is presuppositional, it must be contextually entailed, and contradicting exhaustivity will result in mostly 'false' judgments; see, for example, Abrusán & Szendrői 2013 and Romoli & Schwarz 2015 for results from experimental studies in which there are a majority of 'false' judgments or rejections of presupposition violations.

⁶See, for example, Matthewson 2008 for presuppositional differences between English and St'àt'imcets in determiners and third-person pronouns.
tinue'; proportions for all four sentence types in both experiments are presented in figure 2. Generalized linear mixed-effects models were adopted for the analyses using the *glmer* function of the *lme4* package (GPL-2 | GPL-3: v.1.1-17, Bates et al. 2015b) in the R environment (GPL-2 | GPL-3: v.3.3.7, R Core Team 2017). Treatment contrasts were used: for the factor Sentence, CLEFTS were the baseline comparison for DEFINITE PSEUDOCLEFTS (numeric covariates, with clefts coded as 0, and definite pseudoclefts coded as 1); note that exclusives and focus were not included in the statistical models given floor/ceiling effects making meaningful comparisons difficult. For the factor Language, FRENCH was the baseline comparison for GER-MAN (numeric covariates, with French coded as o and German coded as 1), using the dataset from the identical experiment run on German; see De Veaugh-Geiss et al. 2018. In all models reported, the parsimonious random-effects structures were utilized using the rePCA function in the RePsychLing library (MIT, v.o.o.4)7 following the recommendations made in Bates et al. 2015a.

4 Results

In what follows, we report the French results with direct comparisons to German (from De Veaugh-Geiss et al. 2018).

4.1 Late Response (Box 3/4)

Since the Late Response measures served as a control that participants understood the logic of the experiments, we discuss them first. The logic is as follows.

- In Experiment I: when participants chose 'continue', verifying the canonical meaning was not sufficient to make a truthvalue judgment, and checking that exhaustivity holds must have been relevant enough to motivate further uncovering.
 - \Rightarrow In this case, participants should make a majority of 'true' judgments if exhaustivity holds in Box 3/4 (+EXH), and a

⁷Available at https://github.com/dmbates/RePsychLing.

majority of 'false' judgments if it does not (-EXH).

- In Experiment II, by contrast, when participants chose 'continue', violating exhaustivity was not sufficient to make a 'false' judgment, and thus verifying the canonical meaning must have been relevant enough to motivate further uncovering.
 - ⇒ In this case, we expect participants to make a majority of 'true' judgments if the canonical meaning holds in Box 3/4 (+CAN), and a majority of 'false' judgments if it does not (-CAN).

As can be seen in table 2, these predictions are generally borne out (modulo cases with very few data points; see the denominator of the fractions reported).

| | | | | Exclusive | | Focus | | Def.Pse. | | Cleft | |
|--------|---------------------|--------|---------|-----------|----------|-------|-----------|----------|---------|-------|---------|
| French | Exp. I (Verifier) | [+EXH] | 'true' | 98% (| 122/125) | 100% | o (9/9) | 100% | (96/96) | 100% | (63/63) |
| | | [-EXH] | 'false' | 98% (| 124/127) | 64% | (7/11) | 88% | (74/84) | 68% | (43/63) |
| | Exp. II (Falsifier) | [+CAN] | 'true' | 17% | (2/12) | 99% | (119/120) | 83% | (69/83) | 93% | (85/91) |
| | | [-CAN] | 'false' | 100% | (7/7) | 99% | (118/119) | 96% | (77/80) | 99% | (91/92) |

Table 2 Late Responses as percentages (fractions in parentheses) for [+/- EXH] conditions in Experiment I and [+/-CAN] conditions in Experiment II.

4.2 Early Response (Box 2)

We now concentrate on the Early Responses at Box 2, the critical location for both experiments. For the statistical analysis, we directly compared French to German with the data from De Veaugh-Geiss et al. 2018.⁸ Results are presented graphically in figure 2, in which the observed proportions for the four sentence types are illustrated for French by large solid circles, and for German by large hollow circles. Moreover, back-transformations from the model-estimates for

⁸The models reported here were as follows. Experiment I: $glmer(TVJ.Box2 \sim DefPse*Lang + (1+DefPse | |Participant) + (1+DefPse | |Item), family = binomial). Experiment II: <math>glmer(TVJ.Box2 \sim DefPse*Lang + (1+DefPse | |Participant) + (1 | Item), family = binomial).$

clefts and definite pseudoclefts are illustrated by small solid circles with 95% confidence intervals.



Figure 2 Early Responses ('true'/'false' judgment = 1, continue = 0) at Box 2 for Experiment I (left) and Experiment II (right): presented here are the observed proportions for all sentence types (French = large solid circle, German = large hollow circle) and the back-transformed model estimates for definite pseudoclefts and clefts (small solid circles with 95% CIs).

4.2.1 Experiment I (Verifier)

Looking at the main effect of *Sentence* form in Experiment I, when verifying the canonical meaning at Box 2, French DEFINITE PSEUDO-CLEFTS elicited a significantly lower proportion of 'true' responses compared to the baseline *c'est*-CLEFTS ($\hat{\beta} = -1.9753$, SE = 0.3634, z = -5.436, p = 5.45e-08), suggesting that French participants interpreted definite pseudoclefts with a different strength of exhaustivity than clefts. Turning to the main effect of *Language*, when comparing French *c'est*-clefts (baseline) to German *es*-clefts, there was no significant effect found ($\hat{\beta} = -0.7997$, SE = 0.9652, z = -0.828, p = 0.407407). However, while GERMAN participants treated DEFINITE PSEUDOCLEFTS and CLEFTS on a par, FRENCH participants responded differently to the two sentence forms, with DEFINITE PSEUDOCLEFTS eliciting a lower proportion of 'true' responses than CLEFTS, and this interaction of *Sentence* x *Language* was significant ($\hat{\beta} = 1.6605$, SE = 0.5036, z = 3.297, p = 0.000977).

4.2.2 Experiment II (Falsifier)

Looking at the main effect of *Sentence* form in Experiment II, when encountering a violation of exhaustivity at Box 2, *c'est*-CLEFTS (baseline) and DEFINITE PSEUDOCLEFTS in French showed no statistically significant difference ($\hat{\beta} = 0.5724$, SE = 0.3505, z = 1.633, p = 0.102427). However, turning to the main effect of *Language*, FRENCH clefts (baseline) were significantly less likely to elicit 'false' judgments compared to their GERMAN counterparts ($\hat{\beta} = 1.2962$, SE = 0.6032, z = 2.149, p = 0.031644). Nevertheless, unlike in Experiment I (Verifier), the interaction of *Language* x *Sentence* was not significant ($\hat{\beta} = -0.3881$, SE = 0.4880, z = -0.795, p = 0.426456), with CLEFTS and DEFINITE PSEUDOCLEFTS showing parallel response patterns across the languages tested.

4.3 Post hoc Analysis: Exhaustive vs. Non-Exhaustive Responders

In order to understand better the nature of the intermediate result patterns for cleft constructions, participants were divided into three groups based on their response patterns in this condition: that is, 'exhaustive', 'non-exhaustive', and 'chance' responders. The 'exhaustive' group includes participants who (i) in Experiment I chose at Box 2 to make a 'true' judgment 40% or less of the time (i.e., they instead chose to check that exhaustivity holds by uncovering Box 3/4); and who (ii) in Experiment II chose at Box 2 to make a 'false' judgment 60% or more of the time (i.e., for these participants, the violation of exhaustivity was sufficient to make a 'true' judgment 60% or more of the time (i.e., for these participants who (i) in Experiment I chose at Box 2 to make a 'true' judgment 60% or more of the time (i.e., for these participants who (i) in Experiment I chose at Box 2 to make a 'true' judgment 60% or more of the time (i.e., for these participants who (i) in Experiment I chose at Box 2 to make a 'true' judgment 60% or more of the time (i.e., for these participants who (i) in Experiment I chose at Box 2 to make a 'true' judgment 60% or more of the time (i.e., for these participants, verifying the canonical meaning of the cleft was sufficient to make a 'false' judgment 40% (ii) in Experiment II chose at Box 2 to make a 'false' judgment 40%



Figure 3 Observed proportions for definite pseudoclefts and clefts (French = solid circle, German = hollow circle) for Early Responses ('true' judgment = 1, continue = 0) for Experiment I (Verifier) divided into two participant groups: non-exhaustive responders (left) and exhaustive responders (right).

or less of the time (i.e., for these participants, falsifying exhaustivity was not sufficient to make a truth-value judgment, and instead they chose to continue uncovering Box 3/4). Participants who fell between 40–60% judgment in either experiment were put into the 'chance' group (Experiment I: French 4, German 0; Experiment II: French 1, German 2). (Note that we only report the observed proportions per responder type in this section due to the low number of data points in some conditions.)

Based on this division, we observe the following: broadly speaking, for participants who interpreted clefts exhaustively, this interpretation was very strong in both languages; and vice versa for those who interpreted clefts non-exhaustively (illustrated in the near ceiling and floor effects per responder type in figures 3–4). Moreover, definite pseudoclefts generally patterned with clefts in both languages, with the exception of the non-exhaustive responders in the French version of Experiment I, for whom definite pseudoclefts elicited nearly 50-50 responses (see the left graph in figure 3). Cru-



Figure 4 Observed proportions for definite pseudoclefts and clefts (French = solid circle, German = hollow circle) for Early Responses ('false' judgment = 1, continue = 0) for Experiment II (Falsifier) divided into two participant groups: non-exhaustive responders (left) and exhaustive responders (right).

cially, in terms of the total number of participants for each group:

- In Experiment I, there are two groups splitting participants roughly in half for both French (14 non-exhaustive, 14 exhaustive) and German (13 non-exhaustive, 19 exhaustive), as seen in figure 3.
- In Experiment II, however, the two languages differ: whereas for German the responder-type division was again just about 50-50 (16 non-exhaustive, 14 exhaustive), for French 24 out of the 32 participants fell into the non-exhaustive group and only 7 participants fell into the exhaustive group. Furthermore, definite pseudoclefts patterned on a par.

4.4 Summary of Results

To sum up the results reported above, we draw the reader's attention to the following points:

(i) For French and German, clefts and definite pseudoclefts elicited intermediate response patterns with high variation (cf. the con-

trol conditions for exclusives and focus constructions, which generally showed floor and ceiling effects).

- French vs. German: In Experiment I (Verifier), no significant difference was found between French and German clefts; by contrast, in Experiment II (Falsifier), French clefts showed significantly weaker exhaustivity effects than their German counterparts.
- Clefts vs. definite pseudoclefts: In Experiment I (Verifier), French definite pseudoclefts showed stronger exhaustivity effects than *c'est*-clefts; however, in Experiment II (Falsifier) this effect was not found.
- (ii) In a post hoc analysis, participants generally split into two groups in both languages: those who reliably treated clefts exhaustively, and those who reliably treated clefts non-exhaustively.
 - Exhaustive vs. Non-Exhaustive Responders: while for German Experiments I (Verifier) & II (Falsifier) and for French Experiment I (Verifier) this division roughly split participants in half, for French Experiment II (Falsifier) a clear majority of participants fell into the non-exhaustive group.

5 Discussion and Proposal

Returning to the first research question formulated in section 3.1, we were interested in the systematicity and robustness of exhaustivity in French clefts, in particular compared to their German counterparts. Despite French clefts being noted to associate with exhaustivity (Lambrecht 2001, Destruel 2013), there were reasons to expect that exhaustivity effects might be different, namely weaker, because of their broader discourse semantics and higher frequency of use (see section 2.2).

Our results are, at least in part, compatible with this line of thinking. Although the results from Experiment I (Verifier) do not show a significant difference between the behavior of German and French clefts, we think the more revealing results to consider are those from Experiment II (Falsifier). In this experiment, in which participants were directly confronted with violations of exhaustivity, we found considerable variation in the strength of the exhaustive inference in clefts: in fact, there was a significant difference between the French *c'est*-cleft and the German *es*-cleft, with the response patterns for French participants compatible with a relatively weak exhaustive interpretation. Furthermore, descriptively speaking, there was a lower number of participants falling into the exhaustive group in French compared to German.

None of the theories on cleft exhaustivity discussed in section 2.1 can account for these differences: the fact that exhaustivity was neither robust nor systematic across experiments and participants does not align with semantic theories that predict a hardwired exhaustivity inference, and the variation across languages appear at odds with claims that exhaustivity arises via universal pragmatic principles. So how should one account for these differences?

One might be inclined to argue French is different enough that it should be excluded from accounts trying to model exhaustivity in clefts. We, however, think it is preferable to seek a unified account that can embrace variation across languages and speakers. To this end, we take an approach following De Veaugh-Geiss et al. (2018) (which itself is based on Pollard & Yasavul 2014), in which exhaustivity is neither semantic nor pragmatic, but derived indirectly via the resolution of the anaphoric existence presupposition in clefts (echoing Horn 1981). The crux of our proposal is in the cross-linguistic differences in the derivation of this existence inference.

In a nutshell, exhaustivity arises in how the antecedent of the existence presupposition of the cleft is accommodated (Pollard & Yasavul 2014, De Veaugh-Geiss et al. 2018). If the existence presupposition is accommodated to the maximal discourse referent that answers the Question Under Discussion (QUD), then an exhaustive interpretation arises; if it is accommodated to a non-maximal discourse referent, than a non-exhaustive interpretation arises (see De Veaugh-

Geiss et al. 2018).

Regarding the existence presupposition, Abrusán (2016) claims that for English *it*-clefts this presupposition is derived from the background question generated by the cleft plus the presuppositional constraint that the disjunction of the Hamblin set is true. The QUD of clefts is of the form Who P? (or a sub-question of this QUD), which is derivable directly from the cleft relative. For French, however, we follow Destruel & De Veaugh-Geiss (2018) and claim that the QUD for c'est-clefts-unlike for English or German-can, but crucially need not, correspond to the cleft relative, since it can also signal broad focus; see (4). Given this less strict question-answer congruence, the background question that the cleft answers is potentially ambiguous—especially when little context is provided, such as in our experiment where clefts appeared out-of-the-blue-and the existence inference of French clefts may not arise (see the discussion on hard vs. soft existence in clefts and focus in Abrusán 2016). Consequently, this 'soft' existence presupposition may hinder the derivation of the exhaustive inference, making it a weaker inference.

One advantage of this approach is that differences observed for French do not depend on the nature of exhaustivity being pragmatic or semantic; rather, it is due to the way clefts interact with context and discourse, specifically in terms of the QUD.

Turning to the second research question formulated in section 3.1, we asked how the interpretation of French clefts compared to that of other sentence forms, in particular, definite pseudoclefts. This question was motivated both by several theoretical accounts which treat the two sentence types in parallel ways, as discussed in section 2.1, as well as by the results reported for German in De Veaugh-Geiss et al. 2018. In fact, while the results for German are compatible with a fully parallel analysis of clefts and definite pseudoclefts (albeit not as had been previously proposed in literature; see De Veaugh-Geiss et al. 2018 for details), the results for French are less clear. Indeed, one surprising result was the inconsistencies for French definite pseudoclefts for the section of the section.

definite pseudoclefts compared to *c'est*-clefts in Experiment I (Verifier) which was absent in Experiment II (Falsifier).

Although lacking a full-blown analysis of definite pseudoclefts here, we nevertheless wish to discuss a few observations of interest. First, the differences found between clefts and definite pseudoclefts may in part be compatible with the analysis for clefts presented above. If one assumes anaphoric reference is part of the meaning of French definite pseudoclefts, the derivation of exhaustivity may arguably be similar to that of clefts (and in line with De Veaugh-Geiss et al. 2018 for German), with one critical difference: the discourse semantics of the two sentence forms differ. In fact, based on nativespeaker intuitions, French definite pseudoclefts are not suitable as answers to broad focus questions such as *What happened?* Just as for the argumentation regarding French vs. German clefts above, the stricter question-answer congruence for definite pseudoclefts could explain why exhaustivity was found to be stronger in definite pseudoclefts than in clefts in French, at least in part.

A second observation to note is that the French definite pseudocleft form tested here, *la personne*, may be dissimilar to the form tested in German. In De Veaugh-Geiss et al. 2018 it is claimed that the German definite compound tested (*derjenige* 'that one there') is not a run-of-the-mill definite description, specifically in terms of anaphoricity. In fact, we note that there is a more colloquial form of the definite pseudocleft in French, namely, *celui qui* ... 'that.one-DEMONSTRATIVE who ...', which on the surface—being derived from a demonstrative form—may in fact make for a more direct comparison to the anaphoric definite pseudocleft tested in German. At the time of writing we leave such considerations for future research.

6 Conclusion

This paper tested the robustness and systematicity of cleft exhaustivity in French compared to German and to exhaustivity inferences found in other structures, in particular, clefts and definite pseudoclefts. The main finding is that French clefts appear in part to have weaker exhaustivity effects than in German, whereas the comparison with definite pseudoclefts was less straightforward. We sketched a proposal for a unified account of exhaustivity, following previous analyses in arguing that clefts do not conventionally encode exhaustivity; rather, an exhaustive inference may or may not arise depending on how the anaphoric existence presupposition is resolved, and we account for the cross-linguistic differences reported here in terms of the derivation of this existence presupposition. This work constitutes a modest step towards better understanding the exhaustive inference associated with French clefts.

A Target Stimuli

In order to derive the exclusive, focus, and definite pseudocleft conditions, please compare (1) to examples (2a)–(2c).

- (1) C'est Jean qui a mis un pull.'It is Jean who put on a sweater'
- (2) C'est Marc qui a préparé un cocktail.'It is Marc who prepared a cocktail.'
- (3) C'est Charles qui a changé un pneu.'It is Charles who changed a tire.'
- (4) C'est Pierre qui a fait un repas.'It is Pierre who made a meal.'
- (5) C'est Charles qui a imprimé une carte.'It is Charles who printed a card.'
- (6) C'est Pierre qui a caressé un chat. 'It is Pierre who pet a cat.'
- (7) C'est Jean qui a repassé une chemise.'It is Jean who ironed a shirt.'
- (8) C'est Marc qui a récité un poème.'It is Marc who recited a poem.'
- (9) C'est Charles qui a acheté un tapis.'It is Charles who bought a rug.'
- (10) C'est Pierre qui a cuisiné une tarte.'It is Pierre who baked a pie.'

- (11) C'est Jean qui a fermé une fenêtre.'It is Jean who closed a window.'
- (12) C'est Marc qui a planté un cactus. 'It is Marc who planted a cactus.'
- (13) C'est Jean qui a brossé un cheval.'It is Jean who brushed a horse.'
- (14) C'est Marc qui a tricoté une écharpe.'It is Marc who knit a scarf.'
- (15) C'est Charles qui a porté une échelle.'It is Charles who carried a ladder.'
- (16) C'est Pierre qui a raconté un mensonge.'It is Pierre who told a lie.'
- (17) C'est Jean qui a arrosé une plante.'It is Jean who watered a plant.'
- (18) C'est Marc qui a écrit une lettre.'It is Marc who wrote a letter.'
- (19) C'est Charles qui a lancé une balle.'It is Charles who threw a ball.'
- (20) C'est Pierre qui a escaladé une montagne.'It is Pierre who climbed a moutain.'

E. Destruel & J. P. De Veaugh-Geiss

- (21) C'est Charles qui a vendu un ordinateur. 'It is Charles who sold a computer.'
- (22) C'est Pierre qui a pressé une orange.'It is Pierre who squeezed an orange.'
- (23) C'est Jean qui a lu un livre.'It is Jean who read a book.'
- (24) C'est Marc qui a organisé une fête.'It is Marc who organized a party.'
- (25) C'est Charles qui a dessiné un arbre.'It is Charles who drew a tree.'
- (26) C'est Pierre qui a volé un crayon.'It is Pierre who stole a pen.'

- (27) C'est Jean qui a lavé une assiette.'It is Jean who washed a plate.'
- (28) C'est Marc qui a allumé une bougie.'It is Marc who burnt a candle.'
- (29) C'est Jean qui a bu une boisson.'It is Jean who drank a soda.'
- (30) C'est Marc qui a claqué une porte.'It is Marc who shut a door.'
- (31) C'est Charles qui a épluché une carotte.'It is Charles who peeled a carrot.'
- (32) C'est Pierre qui a regardé un film.'It is Pierre who watched a movie.'

Acknowledgments Authors are listed in alphabetical order. We thank the editor, Christopher Pinon, as well as the anonymous reviewer for helpful feedback on previous versions of this manuscript. This paper has also benefited from discussions with many people, including David Beaver, Edgar Onea, Agata Renans, and Malte Zimmermann, as well as the members of the audience at the CSSP 2017 Conference. This research was funded in part by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—Projektnummer 317633480—SFB 1287, Project Co2, as well as the "Exhaustivity in *it*-Clefts" project in the XPrag.de priority program (SPP 1727). All errors are our own.

References

- Abrusán, M. (2016). Presupposition cancellation: Explaining the 'soft-hard' trigger distinction. *Natural Language Semantics*, 24, 165–202.
- Abrusán, M., & Szendrői, K. (2013). Experimenting with the king of France: Topics, verifiability and definite descriptions. *Semantics & Pragmatics*, 6(10), 1–43.
- Atlas, J., & Levinson, S. (1981). *It*-clefts, informativeness and logical form: Radical pragmatics (revised standard version). In P. Cole (Ed.) *Radical Pragmatics*, (pp. 1–62). New York: Academic Press.

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015a). Parsimonious

mixed models. *Retrieved from http://arxiv.org/abs/1506.04967* (*ArXiv e-print; submitted to* Journal of Memory and Language), 2.

- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015b). Fitting linear mixedeffects models using lme4. *Journal of Statistical Software*, 67(1), 1–48.
- Büring, D., & Križ, M. (2013). It's that, and that's it! Exhaustivity and homogeneity presuppositions in clefts (and definites). *Semantics and Pragmatics*, 6(6), 1–29.
- Carter-Thomas, S. (2009). The French *c'est*-cleft: Functional and formal motivations. In D. Banks, S. Eason, & J. Ormrod (Eds.) *La linguistique systémique fonctionnelle et la langue française*, (pp. 127–156). Paris, France: L'Harmattan.
- Clech-Darbon, A., Rebuschi, G., & Rialland, A. (1999). Are there cleft sentences in French? In G. Rebuschi, & L. Tuller (Eds.) *The Grammar of Focus*, (pp. 83–118). Amsterdam: John Benjamins Publishing Company.
- De Cat, C. (2007). *French Dislocation: Interpretation, Syntax and Acquisition*. Oxford University Press.
- De Veaugh-Geiss, J. P., Tönnis, S., Onea, E., & Zimmermann, M. (2017). An experimental investigation of (non-)exhaustivity in clefts. In R. Truswell, C. Cummins, C. Heycock, B. Rabern, & H. Rohde (Eds.) *Proceedings of Sinn und Bedeutung* 21, (pp. 347–365).

URL https://semanticsarchive.net/Archive/DRjNjViN/

- De Veaugh-Geiss, J. P., Tönnis, S., Onea, E., & Zimmermann, M. (2018). That's not quite it: An experimental investigation of (non-)exhaustivity in clefts. *Semantics and Pragmatics*, 11(3). Early access version. URL http://semprag.org/article/view/sp.11.3
- De Veaugh-Geiss, J. P., Zimmermann, M., Onea, E., & Boell, A.-C. (2015). Contradicting (not-)at-issueness in exclusives and clefts: An empirical study. In S. D'Antonio, M. Moroney, & C. R. Little (Eds.) *Proceedings of the 25th Semantics and Linguistic Theory (SALT 25) Conference*, (pp. 373– 393). First author published under: DeVeaugh-Geiss.

URL http://dx.doi.org/10.3765/salt.v25i0.3054

- Destruel, E. (2012). The French *c'est*-cleft: An empirical study on its meaning and use. In C. Piñón (Ed.) *Empirical Issues in Syntax and Semantics 9*, (pp. 95–112). Paris, France: CSSP.
- Destruel, E. (2013). *An Empirical Study on the Meaning and Use of the French* c'est-*Cleft*. Ph.D. thesis, University of Texas, Austin, Austin, Texas.

- Destruel, E., & De Veaugh-Geiss, J. P. (2018). On the interpretation and processing of exhaustivity: Evidence of variation in English and French clefts. *Journal of Pragmatics*, 138, 1–16. Second author published under: DeVeaugh-Geiss.
- Destruel, E., Velleman, D., Onea, E., Bumford, D., Xue, J., & Beaver, D. (2015). A cross-linguistic study of the non-at-issueness of exhaustive inferences. In F. Schwarz (Ed.) *Experimental Perspectives on Presuppositions*, (pp. 135–156). Springer.
- Dufter, A. (2009). Clefting and discourse organization: Comparing Germanic and Romance. In A. Dufter, & D. Jacob (Eds.) *Focus and Background in Romance Languages (= Studies in Language Companion Series* 112), (pp. 83–121). Amsterdam: John Benjamins Publishing Company.
- É. Kiss, K. (1998). Identificational focus versus information focus. *Language*, 74(2), 245–273.
- Féry, C. (2013). Focus as prosodic alignment. *Natural Language and Linguistic Theory*, 31, 683–734.
- Hamblin, C. L. (1973). Questions in Montague English. *Foundations of Language*, 10, 41–53.
- Hamlaoui, F. (2009). *La focalisation à l'interface de la syntaxe et de la phonologie: Le cas du français dans une perspective typologique*. Ph.D. thesis, Paris 3 Sorbonne Nouvelle, Paris, France.
- Hedberg, N. (2013). Multiple focus and cleft sentences. In K. Hartmann, & T. Veenstra (Eds.) *Cleft Structures*, (pp. 227–250). Amsterdam: John Benjamins Publishing Company.
- Heizmann, T. (2007). Children's acquisition of exhaustivity in clefts. In H. Caunt-Nulton, S. Kulatilake, & I.-H. Woo (Eds.) *Proceedings of the 31st Boston University Conference on Language Development*, (pp. 298–309). Cascadilla Press.
- Heizmann, T. (2012). Exhaustivity in Questions & Clefts and the Quantifier Connection: A Study in German and English. Ph.D. thesis, University of Massachusetts, MA.
- Horn, L. (1981). Exhaustiveness and the semantics of clefts. In V. E. Burke,
 & J. Pustejovsky (Eds.) North Eastern Linguistic Society (NELS), vol. 11,
 (pp. 125–142). Amherst: University of Massachusetts.
- Horn, L. (2014). Information structure and the landscape of (non-)at-issue meaning. In C. Féry, & S. Ishihara (Eds.) *The Oxford Handbook of Infor-*

mation Structure, (pp. 108–127). Oxford University Press.

- Karssenberg, L., & Lahousse, K. (2015). French *il y a & c'est* clefts: A corpusbased analysis of their information structure. Paper presented at Societas Linguistica Europaea 2015 in Leiden, The Netherlands.
- Katz, S. (1997). The Syntactic and Pragmatic Properties of the c'est-Cleft Construction. Ph.D. thesis, University of Texas at Austin, Austin, Texas. Retrieved from ProQuest Dissertations and Theses. (Accession Order No. AAT 9802920).
- Katz, S. (2000). Categories of *c'est*-cleft constructions. *Canadian Journal of Linguistics*, (pp. 253–273).
- Križ, M. (2017). Referentiality, exhaustivity, and trivalence in *it*-clefts. Manuscript, Institut Jean-Nicod.
- Lambrecht, K. (1994). *Information Structure and Sentence Form: Topic, Focus and Mental Representations of Discourse Referents*. Cambridge University Press.
- Lambrecht, K. (2001). A framework for the analysis of cleft constructions. *Linguistics*, *39*, 463–516.
- Matthewson, L. (2008). Pronouns, presuppositions, and semantic variation. In T. Friedman, & S. Ito (Eds.) *Proceedings of the 18th Semantics And Linguistic Theory (SALT 18) Conference*, (pp. 527–550).
- Onea, E., & Beaver, D. (2009). Hungarian focus is not exhausted. In E. Cormany, S. Ito, & D. Lutz (Eds.) *Proceedings of the 19th Semantics And Linguistic Theory (SALT 19) Conference*, (pp. 342–359).
- Pavlou, N. (2015). Explicit and implicit exhaustivity in focus. In C. Brown,
 Q. Gu, C. Loos, J. Mielens, & G. Neveu (Eds.) *Proceedings of the 15th Texas Linguistic Society Conference, Austin, TX*, (pp. 130–149).
- Percus, O. (1997). Prying open the cleft. In K. Kusumoto (Ed.) North Eastern Linguistic Society (NELS), vol. 27, (pp. 337–351).
- Pollard, C., & Yasavul, M. (2014). Anaphoric *it*-clefts: The myth of exhaustivity. In *Proceedings of the 50th Annual Conference of the Chicago Linguistic Society*, vol. 50.
- R Core Team (2017). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/
- Romoli, J., & Schwarz, F. (2015). An experimental comparison between presupposition and indirect scalar implicatures. In F. Schwarz (Ed.)

Experimental Perspectives on Presuppositions, (pp. 215–240). Springer. Shinners, P. (2011). Pygame. http://pygame.org/.

- Tieu, L., & Križ, M. (2017). Connecting the exhaustivity of clefts and the homogeneity of plural definite descriptions in acquisition. In M. LaMendola, & J. Scott (Eds.) *Proceedings of the 41st Annual Boston University Conference on Language Development*, (pp. 651–664). Cascadilla Press.
- Velleman, D. B., Beaver, D., Destruel, E., Bumford, D., Onea, E., & Coppock, L. (2012). *It*-clefts are IT (Inquiry Terminating) constructions. In A. Chereches (Ed.) *Proceedings of the 22nd Semantics and Linguistic Theory (SALT 22) Conference*, (pp. 441–460).
- Xue, J., & Onea, E. (2011). Correlation between presupposition projection and at-issueness: An empirical study. *Proceedings of the ESSLLI 2011 "Workshop on Projective Meaning"*, (pp. 171–184).

Evidentials and Questions

Regine Eckardt • Andrea Beltrama

Abstract The paper investigates inferential evidentials in questions, specifically, German evidential *wohl* and the Italian evidential future. German questions with *wohl* show the interrogative flip. In verb-final syntax, they are interpreted as conjectural questions. We propose an analysis of evidentials in questions based on von Fintel & Gillies's (2011) anchoring of epistemic *might*. The account predicts the interrogative flip and the conjectural reading of verb-final questions with *wohl*. It is extended to Italian questions with evidential future, which convey conjectural questions with-out the interrogative flip. In the final part, we hypothesize how the range of interpretations of evidentials in questions – as found in a wide range of languages – emerges.

Keywords evidential \cdot conjectural question \cdot interrogative flip \cdot future as evidential \cdot German \cdot Italian

R. Eckardt, University of Konstanz, regine.eckardt@uni-konstanz.de

A. Beltrama, LLF/University of Paris 7-Diderot, abeltram@linguist.univ-paris-diderot.fr

In Christopher Pinon (Ed.), Empirical Issues in Syntax and Semantics 12, 121–155. Paris: CSSP. http://www.cssp.cnrs.fr/eiss12/ © 2019 Regine Eckardt & Andrea Beltrama

1 Introduction

Evidential markers have been extensively studied in recent literature (e.g., Faller 2002, Aikhenvald 2004, Davis, Potts, & Speas 2007, Speas 2008). They convey the speaker's type of evidence in support of an assertion *p*. The speaker could claim that Annie sang based on direct perceptual evidence (they heard it), on reportative evidence (others have told him) or on inferential evidence (that Annie sang follows from the speaker's privileged knowledge).

This article investigates evidentials in questions and aims to deepen our understanding of evidentiality at the interface of semantics, pragmatics and syntax. Evidentials in interrogative clauses can trigger two effects, each of which has been claimed to be independent from the other: *interrogative flip*, where the evidential is re-anchored from the speaker to the addressee; and *conjectural questions* (CQs), which express the speaker's curiosity about a certain issue rather than requesting the addressee to answer.

We focus on two case studies: the German inferential evidential particle *wohl* in declaratives, questions and CQs (1a-b); and the use of the evidential future in Italian to mark a question as CQ (2).

- (1) a. Der Schlüssel ist wohl in der Küche. the key is *wohl* in the kitchen 'The key is in the kitchen I assume.'
 - b. Wo ist wohl der Schlüssel?
 where is *wohl* the key 'Where, do you assume, is the key?'
 - c. Wo wohl der Schlüssel ist? where *wohl* the key is 'Where is the key, I wonder.'
- (2) a. La chiave sarà in cucina. the key be.FUT in the.kitchen
 a. 'The key will be in the kitchen.'
 b. 'The key is in the kitchen I guess.'
 - b. Dove sará (mai) la chiave?where be.FUT (ever) the key?'Where (on earth) is the key? (I have no clue.)'

The paper is organized as follows. Section 2 surveys examples of evidentials in questions in different languages, reviewing the accounts that have been put forward to capture their effects. Section 3 reports the readings and contexts of use of *wohl* in German declaratives and questions. Section 4 proposes a two-step analysis for German that covers the flip reading as well as the further CQ reading triggered by verb-final syntax in root questions. Section 5 discusses the Italian *evidential future* in questions and surveys different ways for the speaker to express their pragmatic ulterior motive when asking questions. These motives can be conventionalized as flip questions or conjectural questions. Aligning these options can help us to understand how different languages make use of similar means in different ways. Section 6 summarizes.

2 Background: Two Interpretations of Evidentials in Questions

It has been observed that evidentials in questions give rise to two possible interpretations. One interpretation is the interrogative flip (Aikhenvald 2004, Speas & Tenny 2003, Garrett 2001, Faller 2002) in which the anchor of the evidential shifts from the speaker to the addressee. The Cheyenne hearsay evidential *sėstse* 'I heard that' illustrates this reading (Murray 2009, 2016). An assertion *p* is hedged by *sėstse* to convey '*p*, as I heard'. If *sėstse* occurs in a question *Q* it can be paraphrased as 'given what you heard, what is the answer to *Q*?'. Example (3) shows this for a polar question and (4) for a *wh*-question (Murray 2016).¹

- (3) Mó= ' -é -némene -sėstse Annie?
 Q= EP -3 -sing -RPT.3SG Annie 'Given what you heard, did Annie sing?'
- (4) Tóne'še é-ho'eohtse -sėstse?
 when 3-arrive -RPT.3SG
 'Given what you heard, when did he arrive?'

Another example is the direct evidence marker *te* in Korean as described in Lim 2011. The assertion in (5), marked by *te*, conveys that the speaker has direct evidence for the prejacent.

(5) John-i na-lul po-te-la. John-NOM I-ACC see-te-DECL
'John saw me.'
Implication: The speaker has direct evidence that John saw the speaker. (Lim 2011)

¹RPT = reportative marker, EP = epenthetic segment.

If evidential *te* is used in a question, the speaker requests the addressee to convey eye-witness information in response to the answer.

(6) John-i na-lul po-te-nya? John-NOM I-ACC see-te-Q
'Did John see me?' Implication: The addressee is expected to answer based on their direct evidence relative to whether John saw the speaker or not.

Interrogative flip is crucially not limited to languages with a grammaticized evidential system. English, for example, features this phenomenon with epistemic *might*. Assertions of the form *might S* convey that *S* is possible according to what the speaker knows. When used in a question, *might* instead refers to the epistemic background of the addressee.²

(7) Where might the key be?'What are possible locations of the key, according to what *you* believe?'

Whether a given evidential supports the interrogative flip or not is ultimately encoded as a lexical property of the evidential marker itself. In their survey of evidentials in questions, San Roque et al. (2017) report flip readings for Quiang (Tibeto-Burman), Tsafiki (Barbacoan, South America), Nganasan (Uralic) and Macedonian (Slavic). The present paper argues that German *wohl* poses another example.

The second way to interpret questions with an evidential marker is as CQs (conjectural questions). Other terms used in the literature are deliberative questions, self-addressed questions or questions where no addressee is present. We use this label to refer to questions Q with the following pragmatic profile:³

²We do not intend to claim that *might* is an evidential.

³We avoid the popular criterium "question asked in absence of an addressee" (Jang & Kim 1998, Jang 1999) as this is neither necessary nor sufficient for CQs.

(8) **Conjectural Questions** *Q*:

- a. *Q* conveys the speaker's curiosity about a certain issue
- b. the addressee is not requested to answer
- c. remaining silent is an unmarked reaction for the addressee
- d. the addressee is invited to join the speaker in speculating on the topic at issue
- e. Q roughly means 'I wonder whether Q' where Q' is the truth conditional core of Q'.

The term 'conjectural question' was proposed by Littell, Matthewson, & Peterson (2010), who investigate evidential markers in questions in Státimcets (Lillooet Salish), Nłe?kepmxcín (Thompson Salish) and Gitksan (Tsimshianic). They observe that evidentials in questions lead to an interpretation that fits the profile in (8) (Littell et al. 2010:89). Their following examples from Gitksan illustrate the case.

- (9) Gitksan
 - a. sdin=**ima**=hl xbiist. be.heavy=**infer**=CND box 'The box might be heavy.'
 - b. nee=hl sdin=hl xbiist=a? YNQ=CND be.heavy=CND box=INTERROG 'Is the box heavy?'
 - c. nee=**ima**=hl sdin=hl xbiist=a. YNQ=**infer**=CND be.heavy=CND box=INTERROG 'I wonder if the box is heavy.'
- (10) naa 'an-t gi'nam-(t)=hl xhlaώsxw 'as John?
 who s.REL-3 give-3=CND shirt PREP John
 'Who gave this shirt to John?'
- (11) naa=**ima** 'an-t gi'nam-(t)=hl xhlaósxw 'as John. who=**infer** s.REL-3 give-3=CND shirt PREP John 'I wonder who gave this shirt to John.'

The same reading is also reported in Murray 2016 for Cheyenne evidential *sėstse* in *wh*-questions. The example in (4) has a second reading as a conjectural question, as in (12), in addition to the flip reading reported above.

(12) Tóne'še é-ho'eohtse-sėstse.
when 3-arrive-RPT.3SG
'He arrived sometime, I wonder when.'
(Murray 2016:(53i-ii))

More examples of evidentials giving rise to conjectural questions are listed in San Roque et al. 2017 and the present paper discusses two more cases, German and Italian.

In the extant literature, both the flip reading and the conjecturalquestion reading have been claimed to follow systematically from the interaction between the semantics of evidentials and the semantics of questions.

Lim (2011) derives the interrogative flip as a systematic result of evidential and question semantics. He suggests that the evidential marker *-te* in Korean is combined with the Hamblin denotation of a question, that is, a set of propositions. The evidential combines pointwise with these propositions to yield (proto-)speech acts. All possible answers to the question are thus predicted to mark the respective proposition as direct-witness information of the addressee.

Littell et al. (2010) derive the CQ meaning as a systematic result of evidential and question semantics. They assume that inferential evidentials with prejacent p presuppose 'the speaker has evidence that p'. A question presupposes the conjunction of the presuppositions of all possible answers. Question (11) thus presupposes 'the addressee has inferential evidence that x gave the shirt to John' for all persons x. The authors argue that this presupposition is so strong that the addressee cannot possibly maintain it. Therefore the question is reinterpreted as not requesting an answer, that is, conjectural.

While either analysis captures the relevant data in the respective languages, both proposals seem to suggest that the pattern that they unveil applies crosslinguistically – that is, *all* evidentials in *all* questions in *all* languages should behave in this way. In the remainder of the paper, we argue that this prediction is too strong; however, we remain committed to explaining the emergence of these readings in a principled fashion, and in particular to the idea of linking both the interrogative flip and conjectural readings to the broader semantic properties of questions. Our main case of study is German evidential *wohl* in questions, which shows an intriguing two-fold pattern: it gives rise to the flip reading in questions (Zimmermann 2004, 2008, 2011); it can yield a CQ interpretation when used in root-clause questions with verb-final syntax. Our point of comparison is Italian where the future can be used in an evidential sense and questions in the evidential future are CQ.

3 German Evidential wohl: Data

Let us begin by considering *wohl* in assertions: here, the particle indicates that the speaker has inferential evidence for *p*.

(13) Der Schlüssel ist wohl in der Küche.the key is *wohl* in the kitchen 'The key is in the kitchen I guess.'

For example, the speaker in (13) does not know for a fact that the key is in the kitchen, but they have plausible reasons to believe so: for instance, they might remember that, after returning home, they went in the kitchen to get rid of bags and therefore assume the key is there as well. We'll return to this reading in section 4.1.

In questions *wohl* shows the interrogative flip (Zimmermann 2004, 2011). The question in (14) asks for an answer but at the same time grants permission to the addressee to rely on their inferences and conjectures in addition to knowledge.

(14) Wo ist wohl der Schlüssel?where is *wohl* the key'Where, do you guess, is the key?'

(14) is a true question in that the addressee is requested to provide an answer. Some descriptive grammars diagnose a tendency for (14) to be "self-addressed" (Thurmair 1989), but we assume that their examples are cases of standard questions being contextually coerced into a "self-addressed" stance.

Finally, *wohl* can occur in root questions with verb-final syntax like in (15). V-final questions in German with *wohl* do not request an answer (Lohnstein 2000, 2007, Truckenbrodt 2006, 2013, Altmann 1987). They are CQs in the sense of definition (8) above.

(15) Wo wohl der Schlüssel ist? where *wohl* the key is? 'Where is the key I wonder.'

Earlier authors assume that verb-final syntax alone is the triggering factor for the CQ interpretation (Doherty 1985, Oppenrieder 1989, Altmann 1993, Lohnstein 2000, 2007, Truckenbrodt 2006, 2013, Zimmermann 2013), but this does not seem to apply to *wh*-questions, at the very least. The examples in (16) show that *wh*-questions in verb-final syntax without evidential *wohl* do not have a CQ reading. We use # to indicate this.⁴

(16) a. #Wo der Schlüssel ist? where the key is
b. #Wann der Zug kommt? when the train arrives

⁴Interestingly, evidential *wohl* in verb-final questions can be replaced with evidential *mag* 'might'. The examples in (16) become grammatical as soon as the verb is embedded under *mag* 'might', hence ... *sein mag*, ... *kommen mag*, ... *gesehen haben mag* are acceptable CQs. An example is spelled out in (i).

Wo der Schlüssel sein mag?where the key be may'Where may the key be I wonder?'

German *mag* in this sense is archaic; speakers lack intuitions for declarative sentences. We therefore restrict our attention to *wohl*.

c. #Wen der gesehen hat? who.ACC he seen has

This supports the assumption that evidential *wohl* is a relevant building block for conjectural questions.

Next consider possible reactions to verb-final *wohl*-questions. They fit the profile in (8) above: they do not request an answer, they express an interest of the speaker (Thurmair (1989):144) and are often interpreted as a starter of joint speculations over a given topic, as in (17).

(17) (In a café: A and B observe a deserted mattress leaning against the house. A to B:)
Wer die wohl da hingestellt hat?
Who that-one *wohl* there put has 'Who may have left that one there I wonder?'

What A means to do is engage B in a conversation on the recent history of the mattress. Clearly, B won't be able to provide a straightforward answer to (17) but could have opinions or experience as to what kind of events lead to situations such as the one observed.

Earlier authors propose that verb-final syntax in German is the triggering factor for CQs. They argue that verb-final syntax is typical for subordinate clauses whereas speech acts proper are typically conveyed by main clauses in V2 syntax. They conclude that verb-final questions cannot convey proper questioning acts and thus convey CQs instead, proposing different formal analyses to derive this prediction.

We argue against this view on basis of two observations. First, verb-final syntax is necessary but not sufficient to build CQs. Constituent questions with verb-final syntax but without *wohl/mag* or further particles are not conjectural, as illustrated in (16). Hence an analysis that derives CQs from verb-final syntax fails to explain why further factors are mandatory. Second, contrary to the standard view, verb-final questions can convey proper question acts when they are

used in repeat questions as in the dialogue in (18a-b).

- (18) a. A: Wo ist der Schlüssel? ('Where is the key?')B: [Does not answer]
 - b. A: Hey, wo der Schlüssel ist!? (hey where the key is) 'Hey, do tell me where the key is!?'

Verb-final repeat questions can also be used by the hearer to make sure that they understood the question correctly.⁵

(19) A: Wo ist der Schlüssel? 'Where is the key?'
B: Wo der Schlüssel ist? (where the key is)
(In the kitchen, of course ... / I have no idea ... / Let me think ...)

We therefore conclude that the proposed alignment of verb-final syntax = defective questions (including CQs) as opposed to V-initial/Vsecond syntax = true questions is oversimplified and does not hold.

Before proposing a different analysis of German conjectural *wohl*questions, let us introduce two further data points. First, CQs in German can also be coded by verb-final syntax and *bloß/nur* 'only' which leads to a *wh-on-earth* question.

(20) Wo bloß/nur der Schlüssel ist? where only the key is 'Where in heaven is the key I wonder.'

The particles *bloß/nur* add the exasperated 'can't find the value' (Obenauer 2004, den Dikken & Giannakidou 2002) to standard questions and CQs. They provide a second way to license verb-final syntax and

130

⁵Questions like (18) frequently occur in newspaper texts to suggest that the author is taking up a question of the reader. Given that they are not elliptical in these uses, we propose that verb-final syntax in (18) triggers the back-question interpretation independently of an (elided) antecedent in discourse. Disselkamp (2017) argues on the basis of prosodic evidence that neither conjectural nor back-asking verb-final questions are elliptical.

trigger CQ readings.

Finally, bare polar verb-final questions like (21) are acceptable CQs.

(21) Ob der Schlüssel in der Küche ist?if the key in the kitchen is 'I wonder whether the key is in the kitchen.'

(21) is, however, sensitive to context. Speaker A cannot use (21) as a discourse starter, addressing a bystander B. An introductory *Was meinen Sie?* = 'what's your opinion?' is necessary to clarify the speaker's intention and improves (21) as a discourse starter. CQs with *wohl*, in contrast, can be used discourse-initially. We thus conclude that bare polar verb-final questions like (21) refer to a given topic and are in fact anaphoric (Gutzmann 2011).⁶ While earlier authors take the subtype in (21) as their starting point and consequently ignore the impact of evidentials in CQs, we propose that polar verb-final questions like (21) constitute a case in their own right and disregard them in the following.⁷

We now propose an account of German conjectural *wohl*-questions that rests on two factors: the interrogative flip plus a second pragmatic factor X – coded in German by verb-final syntax – which triggers the conjectural reading. On this view, the analysis of conjectural questions in German conceptually aligns with the accounts proposed for other languages. In addition, in view of the fact that evidential markers figure prominently in conjectural questions of many languages whereas non-standard syntax does not, we think that there is much to gain from extending our focus beyond German to capture the pragmatic factors behind the phenomenon.

⁶This confirms the function of CQs to initiate joined speculations. Alternative phrases like *Ich frage mich* ... 'I wonder' are not suited to improve discourse coherence, which suggests that CQs are not primarily an expression of interest but an invitation. It would be interesting to test this difference in corpora.

⁷Zimmermann (2013) briefly speculates on a concord analysis for *wohl* and verb-final syntax.

4 German wohl in Assertions and Questions

Section 4.1 revisits *wohl* in assertions and proposes a refinement of Zimmermann's earlier analysis. Section 4.2 takes up the challenge of predicting the flip of *wohl* in standard questions and proposes that it rests on general mechanisms of epistemic anchoring in natural language utterances. Finally, we investigate the extra factors that lead to CQs in sections 4.3 and 4.4.

4.1 Wohl in Assertions

Zimmermann (2004, 2011) proposes that *wohl* marks an assertion p as an *assumed belief* of the speaker, which is less reliable than *knowledge*. He uses the predicate ASSUME(A, p) to code that A thinks that p is true but has no certain knowledge.⁸

(22) A: Der Schlüssel ist wohl im Auto. (the key is *wohl* in.the car) Asserted content: p = 'The key is in the car' Non-at-issue: ASSUME(A, p) ASSUME(A, p) is less certain than KNOW(A, p)

Due to the maxim of quantity, *wohl* is restricted to contexts where the speaker assumes but lacks knowledge. Likewise the use of *wohl* in questions is limited to contexts where the speaker believes that the addressee cannot provide a certain answer (Zimmermann 2008).

This analysis, however, faces two challenges. First, Göbel (2018) observes that the following dialogue is felicitous even though the speaker knows the prejacent proposition. Consider the following exchange between A and B: A claims that Rome is in France. B responds: No, it is in Italy, look at this map!

(23) A: Da hab ich mich wohl geirrt. (there have I me *wohl* erred) 'Then I obviously was wrong.'

⁸We deviate from the original analysis according to which ASSUME(A,p) is the one and only asserted content of (22). Zimmermann offers evidence in favor of this claim, yet open issues remain and Murray's (2009) arguments in favor of two-dimensional meaning extend to German *wohl*.

Göbel argues that an analysis in terms of evidentiality can better explain that (23) does not violate the maxim of quantity.

Second, *wohl* cannot be used to mark the prejacent p as highly likely on mere statistical grounds. Imagine a box with 9 white balls and 1 black ball. A knows the distribution, draws a ball out of the box but cannot see the colour. At this point A can assert (24) but (25) would be odd in this situation.

- (24) Ich habe wahrscheinlich eine weiße Kugel gezogen.I have probably a white ball drawn
- (25) #Ich habe wohl eine weiße Kugel gezogen.I have *wohl* a white ball drawn

The use of *wohl* cannot mark a proposition as statistically likely but not certain. For the particle to be felicitous, there need to be specific episodic facts suggesting that A drew a white ball, and which do not reduce to mere probabilistic knowledge. (25) improves, for instance, when A observes that B – who can already see the ball's colour – makes an unsurprised face, which suggests an unsurprising outcome of the experiment.⁹ Such reference to particularly reliable knowledge such as observed episodic facts is reminiscent of von Fintel & Gillies's (2010) notion "privileged knowledge" in their analysis of *must* as marker of logical inferences. Like these we have to leave this notion vague for the moment.

To capture these observations, we propose that a speaker asserting *wohl* p indicates that p is a defeasible inference from their knowledge. The idea can be illustrated by typical uses of *wohl*. Consider a situation where A knows the following:

i. Hein is nowhere to be seen.

⁹We thank Sven Lauer for suggesting this variant of example (25). It is somewhat tricky to delineate "episodic facts" and "general knowledge" here. What counts as "episodic fact" may vary between speakers, even though the contrast (24)/(25) was confirmed robustly by native speakers.

- ii. It is Friday afternoon and Hein usually goes shopping on Fridays.
- iii. His slippers are in the hall.
- iv. The shopping bag is missing.

A can now say:

(26) Hein ist wohl einkaufen gegangen. Hein is *wohl* shopping gone 'Hein has gone shopping I guess.'

Speaker A explicitly grants that new evidence may cause A to retract the inference. If A finds out that Hein's hiking boots are missing as well, A may decide that Hein rather went hunting for mushrooms.

There is more evidence in favour of analysing *wohl* as an inferential evidential. Native speakers of German report the intuition that *wohl* p invites questions in return like *Why do you think so*? or *What makes you believe this*?¹⁰

(27) A: Hein ist wohl einkaufen gegangen.
'Hein has *wohl* gone shopping.'
B: Warum glaubst Du das?
'What makes you think that?'
This reaction suggests interest in A's reasons.

(28) A: Hein ist einkaufen gegangen.B: Warum glaubst Du das?This reaction challenges A's credibility.

While B's question in (27) seems to naturally target the use of *wohl*, the question in (28) somewhat undermines the speaker's authority for the claim, and thus runs the risk of being perceived as offensive.¹¹

¹⁰We thank Ramona Wallner for drawing our attention to this fact.

¹¹The contrast arises most clearly in cooperative question-answer contexts (i.e., outside school exams or lawsuits) and for assertions of non-sensational content. If *wohl* signals that the speaker infers the prejacent p from knowledge, then B's question in (27) is justified as asking for the premisses of the inference.

We spell out the defeasible inference of p in (29), anchoring it to the speaker's epistemic background *Epi*. It is based on defeasible entailment as a logic relation between sets of propositions and propositions (Lewis 1973).¹²

- (29) For individual *A*, let $Epi_{A,w}$ be the set of propositions known by *A* at index *w*. $Epi_{A,w}$ defeasibly entails *p* iff
 - a. there is a finite set of propositions q, q', ...in $Epi_{A,w}$ such that $\{q, q', ...\}$ defeasibly entails p
 - b. there is no additional proposition *r* in $Epi_{A,w}$ such that $\{q,q',\ldots,r\}$ defeasibly entails $\neg p$

We use entails* as shorthand for "defeasibly entails."

For instance, the speaker's inference in (27) is defeasible: the present knowledge entails* that Hein went shopping but additional information could invalidate the entailment*. (30) defines the meaning of *wohl*.

(30) For individual *A*, let *Epi_{A,w}* be the beliefs of *A* at index *w*. *A* utters: *wohl p*↔ *A* conveys
At issue content: *p*Non-at-issue content: *Epi_{A,w}* entails* *p*"My current knowledge entails* that *p*. Further evidence may force me to retract the inference."¹³

¹²There are several ways to implement such a logical relation, all of which are equally suitable for our purposes (Gabbay et al. 1998, Strasser & Antonelli 2016, Reiter 1980).

¹³Defeasible entailment has been extensively studied in artificial intelligence. Formalisms are designed to capture inference patterns that distinguish normal and non-normal cases. For instance, the proposition 'Tweety is a bird' normally allows to infer 'Tweety can fly'. Yet further information can defeat the inference, for instance, the information 'Tweety is a penguin'. Defeasible logic is thus nonmonotonic (more information may mean less inferences) whereas classical logic is monotonic (deductions remain valid even if new information is added).

We write $\langle p \cdot Epi_{A,w}$ entails* p > to distinguish at-issue and non-atissue content, sometimes suppressing the index w.

4.2 Wohl in Questions

We can now build on this analysis to account for the interrogative flip of *wohl* in questions. We take von Fintel & Gillies's (2011) analysis of English epistemic *might* as our starting point. In their proposal, von Fintel & Gillies (henceforth, FG) treat *might* as existential quantifier over the epistemic background of an agent A. We will also say that *might* (like *wohl*) is anchored to an individual A. FG propose that sentences with *might* give rise not just to one denotation – as the semantic composition would have it – but to a set of possible denotations (called a "cloud" by FG). The possible denotations are computed by anchoring *might* to all possible individuals or groups that could play a role in the given utterance situation. To give an example, the sentence in (31), uttered in a context where A talks to B, is assigned the cloud of denotations in (32).

- (31) The key might be in the kitchen.
- (32) $[[might (the key is in the kitchen)]]^{C} = \{ might(Epi_{A}) ('the key is in the kitchen'), might(Epi_{B}) ('the key is in the kitchen'), might(Epi_{A+B}) ('the key is in the kitchen') \}$

 $might(Epi_A)$ ('the key is in the kitchen') is true iff there are worlds compatible with what A knows where the key is in the kitchen. Similarly for *might* anchored to B. Anchored to A+B, *might* quantifies over worlds that are compatible with what A knows and what B knows.

FG propose that the actual denotation under debate is one in the cloud, chosen on basis of general pragmatic principles: a speaker who makes a claim must be authorized to make this claim. In particular, no speaker can make claims about what follows from other speakers' knowledge *unless* the relevant knowledge is known to her. At the beginning of discourse, speakers don't share knowledge beyond world knowledge in the common ground (CG). Thus speaker A is only authorized to use *might* anchored to A. FG illustrate this principle in various types of dialogue such as questions, exam situations, mastermind games, representative assertions and more. If (31) is uttered by A under normal circumstances, the chosen denotation is $might(Epi_A)$ ('the key is in the kitchen').

We propose to generalize the account to *wohl*. (33) recapitulates our denotation for (26) so far.

(33) A: Hein ist wohl einkaufen gegangen. Asserted: 'Hein went shopping' (= p) Non-at-issue: Epi_A entails* 'Hein went shopping' <'Hein went shopping' $\cdot Epi_A$ entails* p>

Assuming FG's analysis of anchored assertions, the content in (33) comes about indirectly. In a first step, the utterance "Hein ist wohl einkaufen gegangen" gives rise to the cloud of possible denotations in (34).

(34) {<'Hein went shopping' (= p) · Epi_A entails* p>, <'Hein went shopping' (= p) · Epi_B entails* p>, <'Hein went shopping' (= p) · Epi_{A+B} entails* p>}

A is authorized to convey that A's knowledge entails* p. A is not authorized to convey that B's knowledge entails* p unless A knows everything that B knows pertaining to the issue whether Hein went shopping or not. Thus A is only authorized to convey $\langle p \cdot Epi_A$ entails* p>, as assumed in (30). The cloud-of-denotations analysis and our earlier, simpler analysis predict the same denotation for declarative sentences, which is empirically adequate.

The indirect account in addition predicts the interrogative flip. We adopt a Hamblin semantics for questions. Following Zimmermann (2008), *wohl* takes scope over the question operator and adds its content to each answer. We thus derive answers with the non-at-issue

element that labels p as a defeasible entailment.¹⁴ The question in (35) gives rise to the cloud of question meanings in (36).

- (35) Wo ist wohl der Schlüssel? where is *wohl* the key
- (36) {{<'key is at z' (= p) · Epi_A entails* p> ; z location}, {<'key is at z' (= p) · Epi_B entails* p> ; z location}, {<'key is at z' (= p) · Epi_{A+B} entails* p> ; z location}}

(36) comprises three sets of answers: 'I have evidence that the key is at z', 'You have evidence that the key is at z', and 'Our pooled knowledge offers evidence that the key is at z'. These represent three question meanings that are abbreviated as Q_A , Q_B , Q_{A+B} in what follows. At the beginning of a discourse the addressee B is not authorized to give answers to Q_A because B cannot know what A can infer about the key's location. B is authorized to answer Q_B .¹⁵ B is not authorized to answer Q_{A+B} because facts known to A might delete B's defeasible inferences. We thus predict that the actual question at issue is Q_B , that is, the interrogative flip for *wohl* in (37).

(37) Wo ist wohl der Schlüssel?'What do you guess where the key might be?'

What remains to be explored are situations in which B happens to know the answer to Q. Defeasible entailment includes classical entailment. The analysis therefore predicts that B can assert known propositions p in response to Q. Due to scalar implicature, answers that B knows for certain should not be labelled with *wohl*. Likewise,

¹⁴We follow Korotkova (2015, 2017) and Matthewson et al. (2007) and assume that evidentials and modals are not necessarily categorically distinct types of expressions.

¹⁵A reviewer suggests that B could attribute defeasible inferences to A, making plausible assumptions about A's knowledge – we could call this "delegated inferencing." Delegated inferencing is possible for *might* (as demonstrated in von Fintel & Gillies (2011)) but not for *wohl*. For reasons of space we can not review the evidence here.

A cannot ask B a *wohl*-question if A believes that B knows the answer for certain. The question in (38) is marked (see Zimmermann 2004, 2008).

(38) #Wie ist wohl Ihr Name? what is *wohl* your name 'What's your name, you guess?'

Zimmermann derives (38) on the basis of Gricean maxims: A can expect B to know the answer to this question for certain. The question in (38) adds extra complexity (*wohl*), while B will use the simpler sentences without *wohl* to answer (scalar implicature). Thus the question is more complex and thus dispreferred in comparison to the question without *wohl* (maxim of manner).

We have analyzed *wohl* as a marker of defeasible entailment, anchored to agent A's knowledge. It is a lexical property of *wohl* that it can trigger clouds of denotations. Our account of the interrogative flip is purely pragmatic which is, we believe, adequate for the phenomenon. The analysis offers an alternative to syntax-based analyses such as Speas & Tenny (2003), where the speaker and addressee are represented as part of the syntactic structure (SpeakerP, HearerP). We maintain that the grammatical status of these phrases as well as their interface to semantics, pragmatics and, finally, the real world, is poorly understood so far. Gärtner & Steinbach (2006) raise further objections against the syntax-based analysis and the present theory offers a viable alternative. We now turn to the second step, which is the derivation of conjectural questions.

4.3 German Verb-Final Questions: wohl and CQs

This section derives German conjectural questions with *wohl* from the cloud $\{Q_A, Q_B, Q_{A+B}\}$. We propose that verb-final syntax has the effect of forcing the denotation Q_{A+B} . In order to see the consequences, we have to spell out in more detail what Q_{A+B} amounts to.

The crucial point is this: If A asserts *wohl p* it refers to A's knowl-

edge. B might know facts that defeat A's inferences. The same holds for B. Therefore, if A and B pool knowledge that pertains to a given issue, they can draw more reliable inferences, inferences that are less in danger of being defeated. Given the nature of defeasible inference, A may have to retract *entailed** beliefs when updated with B's knowledge and vice versa. With $Epi_{A,w}$ = the set of propositions A knows in w, we have $Epi_{A+B,w} = Epi_{A,w} \cup Epi_{B,w}$.

- (39) $Epi_{A+B,w}$ defeasibly entails *p* iff
 - a. there is a finite set of propositions $q, q' \dots$ in $Epi_{A,w} \cup Epi_{B,w}$ such that $\{q, q', \dots\}$ defeasibly entails p
 - b. there is no additional proposition r in $Epi_{A,w} \cup Epi_{B,w}$ such that $\{q, q' \dots r\}$ defeasibly entails $\neg p$

It follows from (39) that defeasible entailment from A and B's pooled knowledge is not the same as defeasible entailment from A and B's CG at the point when Q is uttered. The definition in (39) assumes that A and B first share their knowledge (we assume, realistically, that only propositions that pertain to the issue at hand are relevant) and afterwards draw defeasible inferences. Q_{A+B} therefore asks for better answers than what the knowledge of A or B alone, or their CG before pooling would entail*.

Another consequence of (39) is that the question Q_{A+B} is unanswerable for B.¹⁶ B can infer *p* in answer to *Q* on basis of her own knowledge but she cannot normally anticipate whether A knows facts that challenge the inference. Thus, A cannot rationally request B to answer Q_{A+B} . If B tries to find an answer to Q_{A+B} , she must start by finding out what A knows about the issue and only then guess an answer that their pooled knowledge will support. Another conventionalized reaction for B can be to remain silent: there is no proposition in Q_{A+B} that B is authorized to assert in response to Q, and B even does not have to say it because it follows from the logic of the question. In summary, the possible reactions of B to Q_{A+B} are

¹⁶Unless B happens to know the answer for certain. We discuss this case below.
exactly those that we profiled for CQ in (8) above.

We can now introduce the last ingredient to derive the conjectural reading of German constituent verb-final questions: the silent operator CONJEC, which has three distinctive properties. First, it makes sure that the sentence shows verb-final syntax, second, it is restricted to questions with an evidential; third, it forces the reading that is anchored to a maximum set of speakers. This is captured by the following definition:

(40) CONJEC

- a. **Syntax:** The CONJEC operator is a tacit operator in *C*⁰ of questions. It blocks V-to-C movement and thus ensures V-final syntax.
- b. **Logical and sortal restrictions:** CONJEC is semantically licensed only if the sister node **Q** has a denotation of type <<< s, t >, t >, t >. More specifically, the sister node must be a cloud of questions that arise from different possible anchorings.
- c. **Semantics:** CONJEC(**Q**) maps **Q** to $Q_G \in Q$ that is anchored to the maximal set of interlocutors G. If there are only two salient speakers A and B, CONJEC(**Q**) = Q_{A+B} .

Let us illustrate the effect of CONJEC with an example.

(41) Wo wohl der Schlüssel ist? where *wohl* the key is 'Where is the key I wonder?'

In syntax, V-final questions carry the operator CONJEC in C^0 . This prevents the finite verb from moving to C^0 . The *wh*-constituent is moved to SpecC. CONJEC is a root clause operator that can not occur in an embedded CP.¹⁷ We will briefly consider the case of polar *wohl*-questions at the end of the subsection.

¹⁷We thank the anonymous reviewer for making this clarification.

(42) [Wo_i CONJEC [der Schlüssel wohl t_i ist]_{IP}]_{CP}

We assume that operators take their logical scope before interpretation. The structure to be interpreted is the one in (43). CONJEC has to take highest scope or else the sister denotation will not be of the correct logical type and content.

(43) [CONJEC (wohl (wo der Schlüssel t_i ist?))]

The combination of *wohl* and the question denotation yields a cloud of denotations as in the previous example.

(44) {{<'key is at $z' (= p) \cdot Epi_A$ entails* $p > ; z \text{ location}}, {<'key is at <math>z' (= p) \cdot Epi_B$ entails* $p > ; z \text{ location}}, {<'key is at <math>z' (= p) \cdot Epi_{A+B}$ entails* $p > ; z \text{ location}}}$

CONJEC forces the interpretation that is anchored to A+B.

(45) CONJEC([[wo der Schlüssel wohl ist]])
= {<'key is at
$$z' (= p) \cdot Epi_{A+B}$$
 entails* $p > ; z$ location}

Following (39), the question at issue must be answered by a proposition p of the form 'The key is at location z' such that

- there is a finite set of propositions q, q', ...
 in Epi_{A,w} ∪ Epi_{B,w} such that {q,q',...} defeasibly entails p
- there is no additional proposition *r* in either $Epi_{A,w}$ or $Epi_{B,w}$ such that $\{q, q', \ldots, r\}$ defeasibly entails $\neg p$

The question can be paraphrased as follows:

- Which proposition p = 'the key is at location z' is such that
 - there are propositions q, q', ... that we both know if we pool knowledge and {q,q',...} entails* p
 - and there is no further proposition *r* that **one or the other** of us knows such that {*q*, *q*', ..., ∧ *r*} entails* ¬*p*?

The addressee B has limited ways to react to (45). Unless B hap-

pens to know the answer for certain, B is not authorized to infer an answer because the question requires A and B to pool knowledge. B can either start pooling relevant knowledge with A (i.e., engage in joint speculation) or remain silent, thus confirming A's expectation that she cannot answer the question. The only kind of situation where B can answer questions like (45) is when B knows the answer for certain. If this is the case, B can and will indeed provide an answer.¹⁸

The analysis can be extended to polar verb-final questions with *wohl* if we assume a second lexical entry for the question complementizer *ob* that denotes CONJEC. Like its tacit counterpart CONJEC in constituent questions, conjectural *ob* is restricted to root-clause CPs. The resulting polar CQ invites speculation about the polar question Q.¹⁹

4.4 Finishing Touches

One final observation remains to be captured: the use of conjectural verb-final questions *wohl* Q is restricted to contexts where A believes that B does not know the answer for certain (Zimmermann 2013). We argue that this restriction follows from our analysis. We proceed in two steps. First, we list the possible epistemic situations of addressee B after the verb-final CQ *wohl* Q has been posed. Second, we argue that verb-final *wohl* Q is the optimal choice for speaker A to ask only in situations of the following kind: if A believes that B does not know the answer and neither what A knows about Q.

Our analysis predicts that verb-final *wohl* Q contains the CONJECoperator that forces interpretation as Q_{A+B} . B can be confronted with the request to answer Q_{A+B} in three types of situation:

¹⁸We leave aside uncooperative discourse in lawsuits, games, exams etc.

¹⁹A reviewer raises the issue whether CONJEC should be modelled as a feature or as an operator. A feature-based analysis could unify constituent and polar CQs, but we observe that – unlike predicted by a feature analysis – the complementizer *dass* in verb-final questions is ungrammatical. We have therefore adopted the operator analysis for the time being.

- i. B knows the answer to Q. In this case, B does not need to draw defeasible inferences to find an answer, and likewise does not have to wait for A's knowledge to defeasibly infer an answer. B will provide the answer to Q, which also answers Q_{A+B} .
- ii. B does not know the answer to Q nor what A knows about Q. B is hence not authorized to draw defeasible inferences that rest on pooled knowledge. B can start a conjectural discourse with A or (as another conventionally accepted move) can remain silent or signal consent (e.g., by using the reply *tja*; see Gärtner & Gyuris 2012:417, fn. 45)
- iii. B has gathered A's knowledge about issue Q before verb-final *wohl* Q? is asked. B can therefore answer Q_{A+B} . In this situation answers to Q_{A+B} are the same as to Q_B .

When phrasing the question, speaker A has expectations about what B knows. A must choose the optimal version of question Q, depending on expectations. The plain question Q is shorter and therefore less marked than *wohl* Q in canonical V-second syntax, due to the maxim of manner (Zimmermann 2008). The question *wohl* Q in canonical Verb-second syntax is less marked than the verb-final version in non-canonical syntax, as argued by Thurmair (1989). The speaker should use the least marked version of Q that will suit her purposes.

If A believes that B knows the answer to *Q*, that is, expects situation (i), she will choose the plain *Q* question as the least marked version that suffices to achieve A's goal (Zimmermann 2008).

If A expects situation (iii), she must take into account that *wohl* Q? in verb-final syntax is marked in comparison to *wohl* Q? in canonical V-second syntax (Thurmair 1989). A can ask *wohl* Q? in canonical V-second syntax, which puts out the cloud { Q_A, Q_B, Q_{A+B} } as the general account of anchored utterance meanings predicts in section 4.2. B will use her knowledge to defeasibly derive an answer (marked with *wohl*). This answer is helpful for A: B knows what A knows and possibly more. The comparatively less marked V-second-question

with *wohl* affords A the same answers as the comparatively more marked verb-final question. A will therefore choose the V-secondquestion. Crucially, only in situations of type (ii) can A gain a real benefit from using marked verb-final *wohl* Q?. The operator CONJEC forces the demanding interpretation Q_{A+B} which, under the given circumstances, is not available for the less marked utterances. A's choice of the marked form is justified by the pragmatic benefit provided.

We thus correctly predict that verb-final *wohl* Q? are restricted to contexts in which A believes that B does not know the answer to Q and neither knows what A knows about Q.²⁰

In summary, we have proposed a two-step analysis of German questions with evidential *wohl*. Questions *wohl* Q? in main-clause syntax give rise to a cloud of question meanings of which the one is chosen that B is authorised to answer and that offers the best answers (i.e., those based on maximal knowledge). This predicts the interrogative flip: if the question is posed at the beginning of discourse where B does not know what A believes, B cannot speak on behalf of A.²¹ The interlocutors thus understand that Q_B is at stake. If the question is posed in verb-final syntax, verb-final syntax is triggered by the presence of a CONJEC operator. It forces the reading where Q is anchored to A+B. These questions are doubly marked in comparison to the simple question: they contain an additional particle and show non-canonical syntax. We argued that this restricts their use to situations where markedness is justified by extra pragmatic benefit. This is the case when speaker A believes that addressee B cannot

²⁰For instance, we predict that German CQs with *wohl* are not used after a lengthy discourse where A and B list evidence about *Q* as we could imagine in detective novels where inspectors share information and then try to conjecture the culprit. They could not use *Wer wohl der Täter ist?* 'Who is the culprit I wonder?', and we predict that this is the case.

²¹It should be noted that *wohl*-questions are rarely uttered in the sense 'Tell me what I can guess about *Q*' that was observed for *might* questions in mastermind game situations. The lack of uses as if from the perspective of another is interesting and should be investigated further.

answer Q or shares A's knowledge about Q.

Finally, the link between conjectural questions and joint speculation is an integral part of our analysis. This link has also been reported for other languages. For Shipibo-Konibo, a Panoan language spoken in Peru and Brazil, Valenzuela (2003) says that evidential *mein* in assertions indicates that the speaker is guessing. To illustrate the use of *mein* in questions, she lists questions that are "selfaddressed questions" but also serve to be "engaging in joint discussion" (Valenzuela 2003:(34)). Similar functions were reported about Salish languages (Peterson, pers. comm.) and Italian (Zucchi, pers. comm.) and this function deserves more attention in future research.

The case of German cannot yet explain how evidentials in questions can trigger CQ readings giving rise to the reading triggered by interrogative flip. Therefore our final section takes a closer look at our second example, the Italian evidential future in assertions and questions.

5 Italian Evidential Future in Questions

We discussed how evidential *wohl* in German questions triggers the interrogative flip and conjectural questions. A more frequent pattern, however, is CQs triggered by evidentials without intermediate steps. This raises two important related questions. First, what is the empirical inventory of the possible effects of inserting evidentials in questions? Second, what do such effects share at a semantic or pragmatic level? Exploring these questions in full depth would go far beyond the scope of the current paper; however, we would like to devote the final part of the article to introduce a relevant case study of the phenomenon: the use of future tense and temporal markers in conjectural questions in Italian. After presenting the data, we discuss possible analyses and speculate that CQ readings result from conventionalized ways to react to a question that is marked as "difficult to answer."

5.1 Italian Future as Evidential Marker

Italian possesses a regular way to mark verbs for the future tense. The example in (46) contains the verb 'be' in the future. In its literal meaning, the sentence conveys the statement in (46a) about the future. However, the sentence can also be interpreted in a second sense as in (46b) where the speaker's assertion is marked as uncertain or inferred information.

(46) La chiave sarà in cucina. the key be.FUT in the.kitchena. 'The key will be in the kitchen.'b. 'The key is in the kitchen I guess.'

Mari (2010) argues that the reading in (46b) is available for all Aktionsarten and thus not a last-resort reinterpretation, contrary to what had been claimed in previous literature. She demonstrates that the speaker in (46b) has indirect evidence for his claim, which could be general knowledge or directly observed facts.

Example (47) illustrates the case for general knowledge. A and B are talking about their son's day at school where activities are planned in advance. The dialogue takes place at 11:40, and meals are scheduled for 11:30–12:00.

(47) A: Che cosa farà? (which thing he.do.FUT)
'What might he be doing?'
B: Mangerà. (he.eat.FUT)
'He is eating I assume.'

Another example given in Mari 2010 is in a scenario when A and B hear a noise outside. In answer to A's question, B replies as in (48).

(48) A: Che cosa succede? (what thing happens)B: Arriverà Giovanni. (arrive.FUT Giovanni)'Giovanni might arrive.'

Mari proposes that the "future (in Italian) marks that the speaker has

indirect access to the event" (Mari (2010):(10)). She points out that unlike English *must*, the Italian future is prohibited in cases where the speaker infers a fact by classical logic and world knowledge. English allows the following reasoning: *The ball is in A, B or C. It is neither in A nor B. Hence it must be in C.* In contrast, the Italian evidential future cannot be used in the analogous statement: *La palla è in A o in B o in C. Non è nè in A, nè in B. #Sarà in C.* Mari characterizes this as the "guessing effect" which parallels our observations about German *wohl* as marker of defeasible inferences.

Finally, the Italian future in the evidential reading always takes wide scope with respect to negation, in analogy to *wohl* (Zimmermann 2004, 2008). This is in line with Matthewson's (2015) characteristics for evidentials.

(49) La chiave non sarà in cucina.the key not be.FUT in the.kitchen'I guess that the key isn't in the kitchen.'#'I do not guess that the key is in the kitchen.' (Unavailable)

We take Mari's observations and our own as a starting point in discussing the Italian evidential future (IEF) in questions.

The IEF in questions forces a conjectural "I wonder" reading that does not request an answer. Speakers from northern Italy report the reading for both constituent questions and polar questions.²²

- (50) Dove sarà la chiave?'Where is the key I wonder?'
- (51) Gianni sarà di Amburgo?'Is Gianni from Hamburg I wonder?'

The Italian evidential future in questions patterns with the Salish languages in section 2 in that the examples in (50)/(51) do not show the interrogative flip reading. The next section sketches how eviden-

²²Other varieties might differ slightly. Speakers of Veneto report that polar questions are unacceptable for them while they agree with the judgment for (50).

tials in questions can directly trigger the CQ interpretation.

5.2 A Tentative Analysis

We assume an analysis for IEF in declaratives along the lines of *wohl*, giving rise to a non-at-issue component that the assertion is defeasibly inferred from A's knowledge. As before, A could be the speaker, the addressee or more, depending on the utterance situation (see section 4.1).²³ We, moreover, suggest that the pragmatic profile of CQs in Italian arises in the same manner as we saw in German: speaker A poses the question, assuming that addressee B does not know the answer for certain. The question requests answers based on defeasible inference based on $Epi_{sp(C)+ad(C)}$. If A's assumptions are correct then B cannot answer the question. Again, B can either engage in joint speculation or acknowledge that *Q* is an interesting but difficult question. We propose that the IEF used in context *C* composes with the question denotation as follows.

(52)
$$[\![IEF]]^C \oplus [\![Q]]^C$$

= {< $p_i \cdot Epi_{sp(C)+ad(C)}$ entails* p_i > ; $p_i \in [\![Q]]^C$ }

The parameters sp(C) and ad(C) are directly computed from *C* without detour via a cloud interpretation. This predicts the unavailability of the interrogative flip, as reported by native speakers. The definition in (52) allows the following reactions of B:

- i. If B does not know the answer to *Q* and has not pooled knowledge with A, B cannot answer the question posed. B can engage in speculative discourse or acknowledge the question without answering.
- ii. If B knows the answer for certain, she can and should answer.
- iii. If B does not know the answer but has pooled knowledge about *Q* with A beforehand, B is authorized to answer. B must use the

²³The proposal is tentative in that it competes with the analyses discussed in Mari 2010. We use it to illustrate the principle, leaving it open which version accounts optimally for the Italian data.

IEF or another marker of indirect evidence/uncertainty in her answer.

If speaker A expects situations of type (i), A also expects that B cannot directly answer. With respect to (ii), we observe that the question with the IEF is more complex and therefore marked in comparison to the simple question. This predicts that A will not use the IEF when she believes that B knows the answer for certain (i.e., in type-(ii) situations). Finally, consider situations of type (iii). If A expects that B knows everything that A herself knows about *Q*, we predict that A requests B to produce a defeasible answer. And indeed, Mari's example in (47) suggests that IEF-questions might be possible in situations where the interlocutors want to reconfirm expectations that rest on their shared knowledge. If this is true, then (52) correctly predicts the pragmatics of the IEF in questions.²⁴

From a broader cross-linguistic perspective, the observed behavior of the IEF in questions raise two issues. First, why don't speakers make use of the flip interpretation of questions like (50)/(51) which seems a very logical and undemanding way to make sense of evidentials in questions? And second, is it an accident that questions with inferential evidentials are interpreted as conjectural instead? It seems useful to frame the case in the larger picture of grammaticalization and language change to better understand the dichotomy. A promising first step, in particular, could be to hypothesize that the grammar of an inferential evidential X in language L goes through three stages. In stage 1, the use of X in questions is not licensed.²⁵ In stage 2, speakers become aware of the possible use of X in ques-

²⁴If, however, (47) assumes asymmetric knowledge of A and B, we should add a lexical restriction that prohibits the use of IEF-questions in situations where A and B have maximized their shared knowledge Epi_{A+B} in propositions that pertain to the question. Based on the judgment of one author, there seems to be no indication that this is the case. We defer a more thorough empirical investigation to future research.

 $^{^{25}\}mbox{San}$ Roque et al. (2017) report that the use of evidentials in questions is secondary.

tions Q(X) in the flip interpretation. The use is not yet part of the grammar of the evidential. In stage 3 speakers put this option to use for specific communicative purposes. They could recruit the form Q(X) in order to facilitate answering for the addressee (McCready & Ogata 2007). This reanalysis establishes the flip interpretation for Q(X). Alternatively, they could reinterpret the facilitated question with the implicature "Q is (too) difficult to answer." In this case, the restriction to contexts where A does not expect an answer comes about by pragmatic enrichment, as we often find in grammaticalization. Q(X) thus is interpreted as conjectural question.

If this is on the right track, one could argue that languages like Cheyenne and German verb-second questions exhibit the first developmental path. Languages like Italian, Salish or Shipibo-Konibo provide evidence for the second option. The proposal predicts that languages of the second type should show the interrogative flip in an earlier historical stage. This prediction must be left for future investigation.

6 Conclusion

The paper discussed the connection between evidential markers and questions by considering two frequent interpretations of evidentials in questions: the interrogative flip and CQ. Our two case studies – German *wohl* and the evidential future in Italian – suggest a rather varied picture, in which (seemingly) similar markers give rise to considerably different pragmatic effects. Looking at a broader picture, this variation suggests that analyses for any specific language have to find the right balance between universals and language-specific properties of evidentials. We hope that the current article can represent a profitable starting point for further research in an area that affords intriguing directions for further research.

Acknowledgments We thank Gisela Disselkamp, Donka Farkas, Andreas Haida, Sven Lauer, Floris Roelofson, Aynat Rubinstein and Ramona Wallner, as well as an anonymous reviewer for their valuable feedback. All remaining errors are our own responsibility. Work on this paper was funded

by the DFG Forschergruppe 2111 "Questions at the Interfaces" at Konstanz, which we gratefully acknowledge.

References

Aikhenvald, A. Y. (2004). Evidentiality. Oxford University Press.

- Altmann, H. (1987). Zur Problematik der Konstitution von Satzmodi als Formtypen. In J. Meibauer (Ed.) Satzmodus zwischen Grammatik und Pragmatik, (pp. 22–56). Tübingen: Niemeyer.
- Altmann, H. (1993). Satzmodus. In J. Jacobs, A. von Stechow, W. Sternefeld,
 & T. Vennemann (Eds.) *Syntax: Ein internationales Handbuch*, (pp. 1006–1029). Berlin: de Gruyter.
- Davis, C., Potts, C., & Speas, M. (2007). The pragmatic values of evidential sentences. In T. Friedman, & M. Gibson (Eds.) *Semantics and Linguistic Theory*, vol. 17, (pp. 71–88). New York: Cornell University.
- den Dikken, M., & Giannakidou, A. (2002). From *hell* to polarity: "aggressively non-D-linked" *wh*-phrases as polarity items. *Linguistic Inquiry*, 33(1), 31–61.
- Disselkamp, G. (2017). The pragmatics of speaker repeat questions. Flash talk and poster at *ESSLLI 2017 Student session*. Toulouse, August 17-28, 2017.
- Doherty, M. (1985). Epistemische Bedeutung. Berlin: Akademie-Verlag.
- Faller, M. (2002). *Evidentiality in Cuzco Quechua*. Ph.D. thesis, Stanford University, Stanford, CA.
- Gabbay, D. M., Hogger, C. J., & Robinson, J. A. (1998). *Handbook of Logic in Artificial Intelligence and Logic Programming*, vol. 3: Nonmonotonic Reasoning and Uncertain Reasoning. Oxford University Press.

Garrett, E. J. (2001). *Evidentiality and Assertion in Tibetan*. Ph.D. thesis, University of California, Los Angeles.

- Gutzmann, D. (2011). Ob einer wohl Recht hat? zwei Satzmodustheorien für das Deutsche im vergleich. *Deutsche Sprache*, 39, 65–84.
- Gärtner, H.-M., & Gyuris, B. (2012). Pragmatic markers in Hungarian: Some introductory remarks. *Acta Linguistica Hungarica*, 59, 387–426.
- Gärtner, H.-M., & Steinbach, M. (2006). A sceptical note on the syntax of speech acts and point of view. In P. Brandt, & E. Fuß (Eds.) *Form, Structure, Grammar,* (pp. 213–222). Berlin: Akademie-Verlag.

Göbel, A. (2018). On evidentiality and undirected questions: A new account

of the German particle *wohl*. *University of Pennsylvania Working Papers in Linguistics*, 24(1), 77–86.

- Jang, Y. (1999). Two types of question and existential quantification. *Linguistics*, 37(5), 847–869.
- Jang, Y., & Kim, I. K. (1998). Self-addressed questions and quantifier interpretation. *Korean Linguistics*, *9*(1), 191–209.
- Korotkova, N. (2015). Evidentials in attitudes: do's and don'ts. In E. Csipak, & H. Zeijstra (Eds.) *Proceedings of Sinn und Bedeutung*, vol. 19, (pp. 340–357).
- Korotkova, N. (2017). Diagnosing the semantic status of evidentials. Talk at the workshop *Questioning speech acts*, September 15, 2017, University of Konstanz.
- Lewis, D. (1973). Counterfactuals. Oxford: Blackwell.
- Lim, D. (2011). Evidentials in interrogatives. In I. Reich (Ed.) *Proceedings of Sinn und Bedeutung*, vol. 15, (p. 419–433). Universität des Saarlandes.
- Littell, P., Matthewson, L., & Peterson, T. (2010). On the semantics of conjectural questions. In T. Peterson, & U. Sauerland (Eds.) *Evidence from Evidentials*, vol. 28 of *University of British Columbia Working Papers in Linguistics*, (pp. 89–104).
- Lohnstein, H. (2000). *Satzmodus-kompositionell: Zur Parametrisierung der Modusphrase im Deutschen*. Berlin: Akademie-Verlag.
- Lohnstein, H. (2007). On clause types and sentential force. *Linguistische Berichte*, 209, 63–86.
- Mari, A. (2010). On the evidential nature of the Italian future. *HAL archives-ouvertes.fr*.
- Matthewson, L., Davis, H., & Rullmann, H. (2007). Evidentials as epistemic modals: Evidence from Státimcets. *Linguistic Variation Yearbook*, 7(1), 201–254.
- McCready, E., & Ogata, N. (2007). Evidentiality, modality and probability. *Linguistics and Philosophy*, *30*(2), 147–206.
- Murray, S. E. (2009). Evidentials and questions in Cheyenne. In S. Lima (Ed.) *Proceedings of Semantics of Under-Represented Languages in the Americas*, vol. 5.
- Murray, S. E. (2016). Evidentiality and illocutionary mood in Cheyenne. *International Journal of American Linguistics*, 82(4), 487–517.
- Obenauer, H.-G. (2004). Nonstandard wh-questions and alternative check-

ers in Pagotto. In H. Lohnstein, & S. Trissler (Eds.) *The Syntax and Semantics of the Left Periphery*, (pp. 343–384). Amsterdam: John Benjamins.

- Oppenrieder, W. (1989). Selbständige Verb-Letzt-Sätze. In H. Altmann, A. Batliner, & W. Oppenrieder (Eds.) *Zur Intonation von Modus und Fokus im Deutschen*, (pp. 163–244). Tübingen: Niemeyer.
- Reiter, R. (1980). A logic for default reasoning. *Artificial Intelligence*, 13(1-2), 81–132.
- San Roque, L., Floyd, S., & Norcliffe, E. (2017). Evidentiality and interrogativity. *Lingua*, 186-187, 120–143.
- Speas, P. (2008). On the syntax and semantics of evidentials. *Language and Linguistics Compass*, 2(5), 940–965.
- Speas, P., & Tenny, C. (2003). Configurational properties of point of view roles. In A. Di Sciullo (Ed.) *Asymmetry in Grammar*, (pp. 315–345). Amsterdam: John Benjamins.
- Strasser, C., & Antonelli, G. A. (2016). Non-monotonic logic. In E. N. Zalta (Ed.) The Stanford Encyclopedia of Philosophy. Metaphysics Research Lab, Stanford University. URL = https://plato.stanford.edu/ archives/win2016/entries/logic-nonmonotonic/.
- Thurmair, M. (1989). *Modalpartikeln und ihre Kombinationen*. Tübingen: Niemeyer.
- Truckenbrodt, H. (2006). On the semantic motivation of syntactic verb movement to C in German. *Theoretical Linguistics*, 32(3), 257–306.
- Truckenbrodt, H. (2013). Selbständige Verb-Letzt-Sätze. In J. Meibauer, M. Steinbach, & H. Altmann (Eds.) *Satztypen des Deutschen*, (pp. 232– 246). Berlin: de Gruyter Mouton.
- Valenzuela, P. M. (2003). Evidentiality in Shipibo-Konibo, with a comparative overview of the category in Panoan. In A. Aikhenvald, & R. Dixon (Eds.) *Studies in Evidentiality*, (pp. 33–61). Amsterdam: John Benjamins.
- von Fintel, K., & Gillies, A. S. (2011). *Might* made right. In A. Egan, & B. Weatherson (Eds.) *Epistemic Modality*, (pp. 108–130). Oxford University Press.
- Zimmermann, M. (2004). Zum *wohl*: Diskurspartikeln als Satztypmodifikatoren. *Linguistische Berichte*, 199, 253–286.
- Zimmermann, M. (2008). Discourse particles in the left periphery. In B. Shaer, P. Cook, W. Frey, & C. Maienborn (Eds.) *Dislocated Elements in Discourse: Syntactic, Semantic and Pragmatic Perspectives*, (pp. 200–231).

New York: Routledge.

- Zimmermann, M. (2011). Discourse particles. In K. von Heusinger, C. Maienborn, & P. Portner (Eds.) *Semantics*, Handbücher Sprache und Kommunikation 33.2, (pp. 2012–2038). Berlin: de Gruyter.
- Zimmermann, M. (2013). *Ob*-VL-Interrogativsatz. In J. Meibauer, M. Steinbach, & H. Altmann (Eds.) *Satztypen des Deutschen*, (pp. 84–104). Berlin: de Gruyter Mouton.

Expletive Negation in English, French, and Mandarin: A Semantic and Language Production Model

Yanwei Jin • Jean-Pierre Koenig

Abstract Expletive negation has been documented in many languages, but what explains the similarities in the contexts where expletive negation occurs remains unclear. This paper examines the contexts where expletive negation is found in three languages, English, French, and Mandarin and shows that the contexts where expletive negation is found are the same. A semantic account of the cross-linguistic similarities in expletive negation triggers is provided and a language production mechanism that accounts for the occurrence of expletive negation is outlined: expletive negation arises when the activation of the negation of the trigger's argument proposition is strong enough to be lexicalized.

Keywords expletive negation · language production · French · Mandarin

```
Y. Jin, University at Buffalo, yanweiji@buffalo.edu
J.-P. Koenig, University at Buffalo, jpkoenig@buffalo.edu
```

```
In Christopher Pinon (Ed.), Empirical Issues in Syntax and Semantics 12, 157–186. Paris: CSSP. http://www.cssp.cnrs.fr/eiss12/ © 2019 Yanwei Jin & Jean-Pierre Koenig
```

1 Introduction

Romance languages are famous for their use of what is called in the French linguistic tradition *expletive negation* (henceforth, EN), that is, grammatically licensed uses of a negative marker that does not contribute to the polarity of the proposition which contains it. EN is exemplified in the French sentence in (1) (abbreviations in all our examples follow the Leipzig glossing rules).

(1) J'ai peur qu'il ne pleuve demain.I.have fear that.it NEG rain.sBJV tomorrow'I fear that it will rain tomorrow.'

Sentence (1) consists of a matrix clause and a complement clause. From a logical point of view, what is feared by the speaker is the proposition that it will rain tomorrow, despite the fact that it contains the negative marker *ne* in the complement clause. The optional presence of ne does not seem to matter: whether or not ne occurs in the complement clause, what is feared by the speaker is that it will rain tomorrow. Ne appears to be semantically empty, or at least not to change the polarity of the proposition denoted by the complement clause. EN is attested in other Romance languages (e.g., Del Prete 2008 on Italian, Dindelegan 2013 on Romanian, Espinal 1992 on Spanish and Catalan, Vázquez Molina 2006 on Spanish, Italian and Portuguese) as well as non-Romance languages (e.g., Rubinstein, Sichel & Tsirkin-Sadan 2015 on Hebrew, Inkova 2006 on Russian). A cursory look at the literature suggests many similarities in the contexts where EN occurs. For instance, the predicate FEAR illustrated in (1) (we use small capitals for semantic predicates denoted by words or collocations) triggers EN in many languages, including languages that do not belong to the Indo-European language family (e.g., Shupamem, Niger-Congo, Nchare 2012; Daakaka, Austronesian, von Prince 2012; Basque, Etxepare 2003).

Despite EN uses having been documented in many languages, what explains similarities in the contexts where EN occurs has not received much attention. The French linguistic tradition has provided a near-exhaustive list of contexts where EN occurs in modern and classical French (see Muller 1991 among others) and has provided several semantic accounts of these contexts of occurrence (see Martin 1984 or Muller 2001, among others). But it remains unclear whether the analysis of EN they provide is French-specific. This is particularly so as French is rather unique in having a distinct form for EN, *ne*, rather than the standard negation (*ne*)... *pas* (although *pas* also has EN uses, see Larrivée 1996, and the overwhelming use of *ne* (rather than *pas*) to mark EN is most likely an instance of blocking). In fact, the special form EN typically takes in French may have obscured some cross-linguistic (and possibly universal) tendencies in both the

contexts where EN occurs and the semantic causes of its occurrence.

This paper has two goals: first, establish that the range of contexts where EN is found in French is not unique to French; second, provide an account of the cross-linguistic similarity of EN-triggering contexts. Because of space considerations, we complement existing French data with data from only Mandarin and English, but research we have independently done on Arabic and Zarma-Sorai and consultation of work on dozens of languages suggest our conclusions extend far beyond these three languages. The similarities in the contexts of occurrence of EN across the languages we have looked at suggest that a semantic licensing of EN, as proposed in Martin 1984 or Muller 1991 for French, is on the right track. We propose a model of EN that expands on their work and combines a language production model and a semantic approach to EN whereby triggers are characterized by the fact that they entail that the argumentation proposition and its dual are evoked by the meaning of EN-triggers.

2 Definition and Previous Studies of Expletive Negation

Many linguists use the term *expletive negation* to cover all cases where the presence of a negative marker does not change the polarity of the proposition that contains it (see, among others, Abels 2005, Espinal 2000, Harves 2002, Makri 2013, Portner & Zanuttini 2000, Yoon 2011). Such a view of expletive negation covers biased and rhetorical questions, negative concord, exclamatives, concessives, and polite requests. The definition of EN we assume in this paper covers a more restricted, but semantically more coherent subset of those contexts, namely, contexts where the occurrence of the semantically redundant negation is triggered by the lexical meaning of an operator. Consider (1) again: if we substitute hais '(I) hate' for ai peur '(I) fear', EN cannot occur. EN in (1) thus seems licensed by the meaning of the collocation avoir peur, an hypothesis we pursue in this paper. More specifically, we use the term *expletive negation* to refer to the occurrence of a logically vacuous negative marker that is licensed by the meaning of a verb (or verbal collocation), an adposition, or an adverb that take

a proposition as semantic argument. It is because of the "negative" meaning of triggers, we argue, that a redundant EN marker occurs. We will speak of a triggering (propositional) operator and an argument proposition and will refer to the argument proposition and its dual as p and $\neg p$, respectively. The argument proposition might be syntactically expressed as a non-reduced finite clause, an infinitival clause, a nominalized clause, or a participial clause. (Note that our use of the term *expletive negation* accords with the traditional definition of EN by French grammarians; see Grevisse 1936.)

There are two major approaches to EN in the literature: one focuses on the formal representation of EN at the syntax-semantics interface while the other focuses on what licenses EN. Some authors (e.g., Espinal 1992, van der Wouden 1994) claim that EN is semantically vacuous and explain the occurrence of EN from the fact that certain syntactic structures have negative implicatures (EN is similar in this respect to negative polarity items or negative concord). Other authors (e.g., Abels 2005, Zvoko Dinković & Ilc 2017) treat EN as a real negation but assume it occurs in some unusual clausal position. These authors argue that having EN high up in a syntactic tree also licenses other types of negation, for example, the so-called genitive of negation in Slavic languages. In contrast to these two kinds of syntactic approaches to EN, some authors have focused on the meaning of EN-triggers and try to explain why their meaning leads to the occurrence of EN. Yoon (2011) argues that the occurrence of EN depends on the mood of the embedded proposition while Makri (2013) argues that EN is only licensed in tensed clauses. (A detailed critique of these last two approaches can be found in Zvoko Dinković & Ilc 2017.) Whatever putative pros and cons the approaches we just alluded to may have, they cover only a subset of the possible ENtriggers. In this paper, we discuss triggers not mentioned in any of the extant literature and are hard for them to explain, for example, FORGET. The predicate FORGET can regularly trigger EN in Mandarin and sometimes in French and English (see sections 4 and 5). Since FORGET does not include mood or tense specifications in its complement clause and is not an NPI or negative concord trigger either, it is hard to see how previous analyses could account for the EN-licensing ability of FORGET.

Muller (1991) provides one of the most comprehensive accounts of EN in French. He argues that French EN-triggers are what he calls *opérateurs négatifs inverses*. These triggers have positive semantic cores but can nevertheless be paraphrased with a dependent negation, that is, we have the equivalence X = Y (NEG) or X can be paraphrased "de façon naturelle par Y (NEG)." For instance, the ENtrigger *craindre* 'fear' is not a negative verb and does not mean 'not wish' but can be paraphrased as 'wish (that) NEG.' Likewise, *déconseiller* 'advise against' can be paraphrased as 'advise (that) NEG.' According to Muller, all of the French EN-triggers can be paraphrased in a similar manner. It is the negation that can appear in the paraphrase of the lexical trigger (itself motivated by the trigger's meaning) that surfaces in the complement clause.

We agree with Muller's intuition that the key to explaining the occurrence of EN lies in the meaning of triggers (which Muller's négatif inverse paraphrases is based on). But there are a few difficulties with Muller's approach. First, Muller's approach fails to pay attention to the modality of complement clauses. Some languages (i.e., Mandarin, Zarma-Sorai) employ different expletive negative markers when different triggers are used. Overlooking the modality of the complement clauses caused, we think, the extant literature to fail to realize that French predicates like REGRET or COMPLAIN can trigger EN (see section 4). Second, for several triggers there is no strict equivalence between X and Y (NEG). Muller in those cases uses the symbol \simeq . But what does this near semantic equivalence amount to? The weaker notion of lexical entailment (or strong invited inference) of $\neg p$ in a restricted set of worlds or temporal intervals, as we argue below, is the answer. Finally, Muller's approach is incomplete as it does not explain why EN occurs, why it occurs in very similar contexts in languages that do not use a special negative marker to express EN, or why it is rather rare for a language to have as

broad a range of (grammatically licensed) EN-triggers as French. A semantic account of EN is not enough, as it would predict identical EN uses across languages. This is why our model of EN combines a semantic analysis with a language production model and a distinction between grammatically licensed EN uses and other EN uses. These additions, we believe, explain both the similarity of triggers and differences in grammaticalization of EN across languages.

3 A Near-Exhaustive List of EN-Triggers in French and Mandarin

In the current linguistic literature, only a few studies have tried to present a list of EN-triggering contexts from several languages (e.g., Zvoko Dinković & Ilc 2017 on Slavic languages, Makri 2013 on Hebrew, Greek, Romance, and Slavic languages, Vázquez Molina 2006 on Romance languages, Yoon 2011 on Japanese, Korean, Romance and Slavic languages) or within a single non-Romance language (e.g., Rubinstein, Sichel & Tsirkin-Sadan 2015 on Hebrew, Inkova 2006 on Russian). In order to have a better understanding of the range of EN-triggers and to determine whether the range of triggers found in French is idiosyncratic or may have a deeper semantic explanation, we conducted a near-exhaustive investigation of ENtriggering contexts in several languages, Arabic, French, Mandarin, and Zarma-Sorai. We focus on French and Mandarin in this paper.

Aside from space considerations, the rationale behind our choice of French and Mandarin is that EN has already received quite a bit of attention in both languages. French EN has been studied for a long time and there is a vast literature (e.g., Damourette & Pichon 1911, Le Bidois & Le Bidois 1935, Gaatone 1971, Grevisse 1936). Furthermore, several monographs on French negation (e.g., Muller 1991, Larrivée 2004, Vázquez Molina 2006) have been published which not only provide a comprehensive list of EN-triggers but also provide a historical context for the use of EN. Because of this extensive previous work, we already have a near-exhaustive list of EN-triggers in French that provides a point of comparison. Within Chinese linguistics, researchers have started to notice the existence of EN since the publication of Zhu's (1959) paper on the EN-trigger chà-diănr ('almost, miss a bit'), which is the most frequently cited EN-trigger in the English literature. Over the ensuing years, scholars have discussed more and more triggers. For instance, Lü (1980/1985) showed that yǐqián 'before', xiǎoxīn 'watch out', and nánmiǎn 'inevitable' trigger EN; Zhang (2004) showed that houhui 'regret' and zébèi 'blame' also trigger EN. In the last ten years, Jiang (2008), Yuan (2012), and a few Master's/Ph.D. theses (Cao 2007, Song 2012, Wang 2008, Wang 2012) discuss various kinds of redundant negations in Mandarin including what we call EN in this paper. For the purposes of the research we report on, we added to all the triggers listed in previous studies verbs retrieved from the Mandarin Verbal Usage Dictionary that we found in corpora to license EN. Our study almost doubles the list of Mandarin EN-triggers and suggests that the contexts that trigger EN in Mandarin are basically the same as those found in French.

4 Our Account of EN

We detail in this section the three components of our account of EN. The first component is concerned with the mechanism responsible for the occurrence of EN. Properties of language production explain, we believe, why on certain occasions EN occurs. Simply put, EN-triggers semantically entail (or strongly imply) the negation of their argument proposition and it is this inference that explains the occurrence of EN. According to Dell's (1986) spreading-activation theory of production, what is entailed or can be inferred from the meaning of a sentence may be activated during speech-production planning and lead to speech errors that lexicalize what is being entailed or inferred. The following quote summarizes the critical component of this model for our purposes:

[...] in the planning of an utterance many concepts would legitimately become activated that would not actually appear in the utterance. This background activation might include activation from concepts that were either presuppositions or inferences that were necessary in the semantic and pragmatic planning of the utterance. (Dell 1986:290–291)

The particulars of Dell's model or the many competing models proposed since are not critical. What matters is the idea that semantic features associated with inferences from what is said can interfere with the planning of an utterance and lead speakers to produce something different from what they intended. We suggest that something like what Dell is proposing is the mechanism that underlies the production of EN: a speaker intends to say p, but because $\neg p$ is strongly activated by the meaning of a trigger, $\neg p$ is produced. Furthermore, because *p* and $\neg p$ are typically entailed (but relative to distinct sets of worlds or time intervals, see below) by the meaning of EN-triggers, the likelihood of occurrence of EN is higher than for other kinds of inferences. It is because the production of EN is more frequent than typical speech errors that EN may become grammatical in some languages. Note that Dell's model accounts for the fact that, although the overwhelming majority of triggers entail $\neg p$ (in a restricted set of worlds or time intervals), that proposition seems only strongly implied by a couple of triggers (see the next section for details).

Second, we distinguish between what we call *highly entrenched* EN uses (in the sense of Langacker 1987, i.e., EN uses which have become "virtually automatic" through repetition) and *low-entrenchment* EN uses. Previous discussions of EN have focused on "standard" language uses. This is particularly true of the discussion of French *ne* and the contexts where it can and cannot occur. This is also true of the absence of EN in English reference grammars or of the mention of EN examples in Horn 2010. But to get to what causes the occurrence of EN, it is critical, we believe, to realize that EN occurs quite often in non-standard language uses and in a wider set of contexts than grammarians recognize. This is particularly true in English, as we discuss in section 5: there are attested examples of EN in English corpora in all the contexts where it occurs in French or Mandarin.

But, it is also true in French: EN is said not to occur in some of the contexts where we found it to occur despite the fact that attested examples can be found in French corpora. To give an example, both Martin (1984) and Muller (1991) state that EN cannot occur in the complement clause of the verb *regretter*. But, as (2) shows, EN *does* occur in this context, although it is relatively rare and thus has a low degree of entrenchment in Langacker's (1987) sense. We call the use of EN in (2) *low-entrenchment* EN.

Je regrette qu'il faille souvent attendre des (2)ne regret that.it NEG should.SBJV often Τ wait INDEF.PL années avant que l'histoire ne juge les tyrans. years before that DEF.history NEG judge.SBJV the tyrants 'I regret that it should often take years before history judges tyrants.' (https://tinyurl.com/2m9r61, accessed on 2018-05-05, article from 2017-12-12)

The reason it is important to consider both highly entrenched and low-entrenchment EN when investigating the semantic conditions under which EN occurs is that what may look idiosyncratic if one only considers a language's highly entrenched EN uses may not look so when both highly entrenched and low-entrenchment EN are examined. When talking about cross-linguistic tendencies in the semantic licensing of EN we thus must be careful to distinguish between highly entrenched EN and EN simpliciter, as more general cross-linguistic tendencies emerge when both kinds of EN uses are included. There seem to be many more idiosyncrasies in the relative entrenchment of EN-triggers than in the occurrence of EN. As an anomyous reviewer points out, our distinction between highly entrenched and low-entrenchment EN raises important issues about the gradient nature of grammaticality and is reminiscent of the work of Lau et al. (2017). We cannot discuss the issue in detail in this paper. Suffice it to say that we view EN in the complement clause of a trigger as starting out as a speech error that may become a more stable property of native speakers' internal grammars in some languages for some

triggers. The language production component of our model of EN explains why such speech errors occur with some frequency across languages. But, as of now, we have nothing of interest to say about why this stabilization (entrenchment) is higher in some languages than others.

The third, semantic, component of our account is inspired by Heim's (1992) work on the semantics of attitude verbs (based on Hintikka 1969 and Kratzer 1981). Simply put, the meaning of attitude verbs entails that their propositional argument is true in a particular set of worlds (e.g., for believe, the set of worlds compatible with the attitude holder's beliefs). The semantic part of our model of EN makes precise the notion of *négatif inverse* paraphrase or near paraphrase discussed in Muller 1991 or the possible worlds approach of Martin 1984 and extends its coverage to the EN-triggers found in Mandarin but not in standard French. We argue that all the contexts that license EN obey one of the following four necessary conditions. We do not attempt to subsume these four conditions under a single condition, as the sentence-production cause of EN means there is no a priori reason to expect triggers to obey a single semantic condition, aside from the fact that the meaning of all EN triggers must entail or strongly imply a negative proposition. (3) should thus be seen as mere generalizations over individual triggers, not as necessary or sufficient licensing conditions in the traditional sense: it is the meaning of each individual trigger that activates the dual of its argument proposition and the production of EN.

- (3) a. **Propositional attitude and speech report triggers:** An operator can trigger the occurrence of an expletive negation in its argument proposition if its meaning entails (or strongly implies) the existence of two distinct propositions of the form $Op_1(p)$ and $Op_2(\neg p)$, where p and $\neg p$ are true in different sets of worlds, as determined by the meaning of Op_1 and Op_2 .
 - b. **Temporal operators triggers:** A temporal operator of the form $Op_1(q, p)$ can trigger the occurrence of an expletive negation in its argument proposition if its meaning entails (or strongly implies)

that $\neg p$ is true at reference time.

- c. **Logical operators triggers:** A modal operator (*impossible*) or propositional functor (*unless*) that includes in its meaning ¬ can trigger the occurrence of an expletive negation in its argument proposition.
- d. Comparative triggers: An operator can trigger the occurrence of an expletive negation in its argument proposition if its meaning entails the existence of two propositions of the form Q(Y, D) and ¬Q(Y, D') (Y possesses property Q to degree D and D' respectively).

We detail how these four generalizations apply to individual triggers in the next section and merely illustrate the first generalization here. French *avoir peur* 'be afraid' in (1) is a propositional attitude trigger that entails both the positive proposition that it will rain tomorrow, which is true in all possible worlds consistent with the speaker's fears (Op_1), and the negative proposition that it will not rain, which is true in all worlds consistent with the speaker's desires or hopes (Op_2).

To summarize, our hypothesis is that EN arises when the meaning of an operator leads to the strong activation of both p (the operator's argument proposition) and its dual $\neg p$ (this formulation must be altered for the complement of comparatives, see (3d)). The strong activation of both propositions stems from the fact that the meaning of triggering operators entails (or strongly imply), to take (3a) as an example, both $Op_1(p)$ and $Op_2(\neg p)$. The semantically conditioned activation of $\neg p$ alongside p is what sometimes leads speakers to produce a logically vacuous negation.

Our model predicts that even in those languages where EN has not been claimed to exist, it can still occur in the form of the unintended expression of an entailed negation during the process of language planning. It also predicts that the meaning of EN-triggers need not be the sole cause of the occurrence of EN, since there can be multiple factors that influence the activation of $\neg p$. For example, as is well-known in the French literature, some triggers must be negated or questioned for an EN to occur; as we suggest somewhat speculatively in section 7, the occurrence of a negation (or a question) in the matrix clause may boost the activation of $\neg p$ and thus increase the likelihood of the expression of $\neg p$. Similarly, since degree of activation is a continuous measure, some triggers might be stronger than others, potentially explaining differences in triggering potential across languages or among triggers with somewhat similar meanings within a language. Although we will not have much to say about this issue in this paper, our model has the ability to account for the graded aspects of EN occurrence.

Finally, a consequence of our three-pronged model of EN is that we need to distinguish between a universal component and a languagespecific component of EN. The universal component is the fact that, in the planning of an utterance, if the semantically conditioned activation of $\neg p$ is sufficiently strong (for a certain trigger), EN may be selected when the speaker lexicalizes her message. The languagespecific component is the degree of entrenchment or grammaticalization of EN after some triggers. The universal potential for the occurrence of EN is caused by the meaning of the triggering contexts and general planning processes of speech production. As long as both $\neg p$ and p are evoked by a context, EN should be attested (modulo syntactic idiosyncrasies, for example, the syntactic expression of the operator's argument). But how frequently EN is attested after a particular trigger in a particular language is a matter of entrenchment and something that is language specific and for which we do not provide an explanation in this paper. Whether EN has a special form and normally only occurs in the context of EN-triggers, for example, the special negative marker ne in French or the negative complementizer ne/quin in Latin, is also language specific (and quite rare, as far as we can see) and depends on whether the expression of the semantically conditioned activation of $\neg p$ has grammaticalized: generally, lexical items that realize the activation of $\neg p$ are ordinary clausal negative markers in the target language, for example, Mandarin bù (imperfective negative marker)/méi (perfective negative marker)/bié

(imperative negative marker)/bù-gāi ('shouldn't').

5 A Semantic Analysis of EN-Triggers in French and Mandarin

In this section, we offer a detailed semantic analysis of each kind of EN-triggers we found in French and Mandarin. We group triggers by the hypothesized entailments that license the occurrence of EN. As in any classification, some triggers can be part of multiple classes and nothing substantial hinges on our particular choice of classificatory scheme. Most of the French examples in this section, unless otherwise specified, were constructed; but they follow the frequently cited examples found in the French literature. The Mandarin examples, with only one exception, are all taken from either CCL (Corpus for Chinese Linguistics, http://ccl.pku.edu.cn) or other web sources. This is because Mandarin EN remains unfamiliar to most non-Chinese linguists and the extant English literature only documents ALMOST and BEFORE as EN-triggers.

Want(X, $\neg p$) Triggering verbs in this class entail that the attitude holder (abbreviated as X hereafter) wants or hopes the negation of the argument proposition to be true. (4) and (5) are partial lists of such verbs in French and Mandarin (° below and throughout this section indicates that the verb strongly prefers to be negated, or questioned, for EN to occur in its complement clause). The meaning of verbs like French *fear* or Mandarin *pà* 'fear' entails that $\neg p$ is true in all worlds corresponding to the attitude holder's desires. When one uses these predicates, two sets of worlds are thus activated. One set contains all the worlds where X's fears are true and *p* denotes what X fears or is apprehensive about; the other set contains all worlds which correspond to X's desires, where $\neg p$ is true. Given the meaning of these verbs, both *p* and $\neg p$ are activated during sentence production. If the activation of $\neg p$ is too strong or $\neg p$ is not inhibited enough, a redundant negative marker will occur. It is the entailed existence of two sets of worlds evoked by these verbs

that explains, we hypothesize, why EN is found in their complement clause.

- (4) craindre 'fear'; l'angoisse que 'the anxiety that'; l'anxiété que 'the anxiety that'; appréhender 'apprehend'; un/le danger que 'a/the danger that'; l'écueil que 'the danger that'; être effrayé que 'be frightened that'; s'inquiéter que 'worry that'; obsession que 'obsession that'; de peur que 'of fear that'; redouter 'fear'
- (5) pà 'fear'; hàipà 'fear'; kǒngjù 'fear'; dānxīn 'worry'; yōulù 'worry'; chóu 'worry'; jiāolù 'be anxious'; zháojí 'be anxious, worry'

(6) and (7) are typical French and Mandarin examples, respectively. In (7), the negative marker must be *bié*, the imperative negative marker in Mandarin. *Bié* is the only expletive negative marker allowed when FEAR predicates are involved. We suspect that this restriction is due to the fact that the meaning of verbs that denote FEAR involves an evaluation of the argument proposition (its undesirability).

- (6) J'ai peur qu'il ne pleuve demain. (= (1))
 I.have fear that.it NEG rain.SBJV tomorrow
 'I fear that it will rain tomorrow.'
- (7) dài bă săn ba! wǒ pà míngtiān duōlúnduō bié xià-yǔ.
 take CLF umbrella BA I fear tomorrow Toronto NEG fall-rain.
 'Take an umbrella! I fear that it might rain in Toronto tomorrow.' (Produced by the first author's mother)

Ought.to($\neg p$) Triggers in this class have to do with behavioral standards. We take the predicate REGRET as an example. REGRET entails that $\neg p$ is true in all worlds that correspond to X's behavioral standards. In all worlds where X's regrets are true, p is true and denotes the content X regrets; but in all the worlds consistent with X's behavioral standards, $\neg p$ is true. It is again the concurrent activation of both p and $\neg p$ (via inference) that leads to the occurrence of EN.

(8) is a Mandarin example; a partial list of Mandarin verbs in this class is provided in (9). The Mandarin expletive negative marker in (8) means something like *shouldn't*.

- Chén Yìpíng lěngjìng (8) chéngmò yǐjīng shĭ le xůduō, tā already make PN calm.down PFv a.lot he silence shènzhì hòuhuǐ zìjǐ **bù**-gāi zhèyàng cāngcù, zuìqǐmǎ yīnggāi regret self NEG-should like.this hasty at.least should even gěi Xuē Fēi yìdiǎn shíjiān. zài again give PN a.little.bit time 'Silence had made Yiping Chen calm down a lot. He even regretted that he had been so hasty. At the very least, he should have given Fei Xue a little time.' (CCL)
- (9) huĭ 'regret'; hòuhuĭ 'regret'; bàoyuan 'complain'; zébèi 'blame'; zéguài 'blame'; guài 'blame'; mányuàn 'blame'; pīping 'criticize'; nánguò 'be sad'; bàoqiàn 'be sorry'; °bù-gāi 'shouldn't'

In all of the studies on French EN to date, no verb in this class has been listed as a possible EN-trigger. Muller (1991) even used *regretter* as a counterexample to Martin's (1984) worlds analysis of EN. But, (2) above showed that REGRET can trigger EN in French.

Believe(X, $\neg p$) or say(X, $\neg p$) The meaning of triggers in this class entails that $\neg p$ is true in all worlds corresponding to X's beliefs. Many of the triggers in this class are verbs of speech reports, so $\neg p$ is true according to X's belief, if X is sincere or the felicity conditions on the speech act denoted by the trigger are satisfied (see Searle 1969). One verb, French *cacher*, is a verb of speaking and $\neg p$ is not entailed but strongly implicated. In a sentence such as (10) (cited by Muller 1991:373), there is no entailment that the speaker will say $\neg p$.

(10) Nous ne pourrons cacher que Madame Guyon ne soit
 We NEG can.FUT hide that Madam Guyon NEG be.SBJV
 trop protégée.
 too.much protected
 'We will not be able to hide that Madame Guyon is protected too

much.'

Two sets of worlds are again activated by the meaning of triggers in this class. One set only contains worlds that do not conform to X's beliefs, but conforms to somebody else's beliefs and in this set, *p* holds; the other set contains worlds compatible with X's beliefs and it is in this set that $\neg p$ is true. (11) and (12) provide partial lists of triggers in French and Mandarin, respectively. Many verbs can only trigger EN in the complement clause when negated. It is worth mentioning that the trigger DESPAIR (in its use as a propositional attitude verb) entails not only that X believes that $\neg p$, but also that $\neg p$ is true in the real world (or w₀) and that X wants p.

- (11) °contester 'question'; °désavouer 'repudiate'; °nier 'deny'; °disconvenir 'deny'; °s'oppose à 'oppose'; °douter 'doubt'; °douteux que 'doubtful that'; °doute que 'doubt that'; °cacher 'conceal, hide a fact', °désespérer 'despair'
- (12) huáiyí 'doubt, suspect'; zhìyí 'question'; dǐlài 'deny'; fǒurèn 'deny'; yǐnmán 'hide a fact'; °fǎnduì 'oppose'; °bù-chéngrèn 'not admit'; °bùxiāngxìn 'disbelieve', °bú-bào-xī-wàng 'despair'
- (13) Niez-vous qu'il ne soit un grand artiste? deny-you that.he NEG be.sBJV a great artist'Do you deny that he is a great artist?'

Expletive Negation in English, French, and Mandarin

zhíchū zìjí céng shōushòu Zhāng Jīnfèng (14) duìyú vuángào faced.with prosecutor point.out self once receive 8000 yuán hǎochùfèi suðyi cái bāngzhù Zhāng Jīnfèng 8000 yuan commission so only.then help PN chǎofáng yìshuō, Zhào Yì fǒurèn zìjǐ méi jiēshòu real.estate.speculation claim deny self NEG receive \mathbf{PN} Zhāng Jīnfèng yìfēnqián. a.single.cent PN

'Faced with the prosecutor's (Jinfeng Zhang) claim that he (the defendant) once accepted 8000 yuan from Zhang as commission and then helped him with real estate speculation, Yi Zhao (the defendant) denied that he accepted money from Zhang.' (*Beijing Times*, 2013-09-17)

 $\neg p$ in w_0 Triggers in this class all entail that $\neg p$ is true in the real world (or w_0). Consider the predicate FORGET (in the sense of *Barbara forgot to get coffee*). It entails that p is true in all worlds where X did what (s)he was supposed to do, but $\neg p$ is true in the real world X is in. It is the fact that $\neg p$ is true in the real world and p in some other worlds that motivates the appearance of an expletive negative marker. The occurrence of EN in the complement clause of some of the triggers (corresponding to English *forget*, *delay*, *stop*) have not yet been mentioned in the extant French or Mandarin literature. For brevity, we provide one example from French and one example from Mandarin. We provide a partial list of Mandarin triggers, as the range of triggers seems wider in Mandarin. Note that in the case of Mandarin *yánhòu* 'delay' or *tuīchí* 'postpone', $\neg p$ is true at w_0 at reference time, but may be false later on.

(15) Vous avez oublié de ne pas nommer Jacques Stephen Alexis, you have forgotten of NEG NEG nominate PN
un grand des grands savants.
one great of.DEF.PL great savants
'You have forgotten to nominate Jacques Stephen Alexis, one of the greatest savants.' (https://tinyurl.com/yxucz2ul, accessed on 2018-05-05)

"Dǐ lǎoshī" chēng xiànzài chūmén qù gěi lǐngdǎo bàn diǎn (16)go for leader teacher say go.out a.bit now do PN bú bàngōngshì, dàn chūmén shí shì, zài wàngjì **méi** affair NEG exist office but go.out when forget NEG dài giánbāo. bring wallet

'Teacher Di (the swindler) said he was out doing something for the leader, and he was not in the office, but he forgot to bring his wallet when going out.' (*Gansu Daily*, 2018-03-23)

(17) wàng 'forget'; wàngjì 'forget'; yánqī 'delay'; yánhòu 'delay'; yánhòu 'delay'; tuīchí 'postpone'; tuīhòu 'postpone'; tíngzhǐ 'stop'; tíngxià 'stop'; fàngqì 'give up'; jùjué 'refuse'

An interesting subclass of triggers is listed in (18) and (19).

- (18) il s'en faut que 'almost' (lit. 'it is necessary from it'); peu s'en faut que 'almost' (lit. 'little is necessary from it'); pour peu que 'almost' (lit. 'for little that')
- (19) chà-diǎnr 'miss-a-bit, almost'; jīhū 'almost'; xiǎnxiē 'almost'; chà 'to be short of'; shǎo 'to be short of'

Triggers in this subclass denote predicates that entail not only that $\neg p$ is true in the real world w_0 but that p is true in worlds minimally different from w_0 .

 $\neg p$ at Reference Time Triggers in this class are temporal operators. The meaning of *before* (as in *q before p*) entails that when *q* is true at reference time, $\neg p$ is also true. The fact that both *q* and $\neg p$ are true at reference time gives rise to logically vacuous EN markers in many languages, including French and Mandarin. Similarly, the meaning of *since* (as in *it's been some time since p*) entails that at reference time, $\neg p$ is true. One might wonder if *after* (in *q after p* should not also be a trigger, as it seems to be the mirror image of *before*. The answer is *No*, as the meaning of *after* does not entail that $\neg p$ is true at reference time (*Bob left after I arrived* does not entail that I had not arrived at reference time). This class of triggers in French also includes a few

verbs of mental attitude that carry a temporal entailment such as *en attendant (que)* 'waiting (for)', *s'attendre à* 'expect', *tarder à quelqu'un (que)* 'cannot wait (for)'.

- (20) Agissons avant qu'il **ne** soit trop tard. we.act.IMP before that.it NEG be.SBJV too late 'Let's take action before it is too late.'
- (21) méi jiéhūn qián, wǒ wàichū bǐsài zǒngshì xiǎng NEG get.married before I be.out match always miss nǚpéngyou. xiànzai yǒu le háizi, zhǔyào xiǎng háizi. girlfriend now have PFv child mainly miss child 'Before I was married, I always missed my girlfriend when I was out playing matches; now I have kids, and I mainly miss them.' (CCL)

None of the Mandarin literature on EN mentions that SINCE can trigger EN. We found several examples in corpora and provide one below.

(22) nàtiān tā shuō xiàbān yǒu yìngchou, kěshì jùlí wǒmen that.day he say off.work have social.activity but since we méi jiànmiàn yí gè lǐbài le, wǒ nǎo le. NEG meet one CLSFR week PFV I annoy PFV 'That day he said he had a social activity after work, but it had already been a week since we met and I got annoyed.' (https: //tinyurl.com/ajkumog, accessed on 2018-05-05)

q WITHOUT p (as in *he left without me knowing it*) also entails q and $\neg p$ at reference time and EN occurs in French in the complement of this trigger, as expected and shown in (23). The concept WITHOUT is not lexicalized in Mandarin, therefore no EN example is forthcoming.

(23) Je l'ai fait sans qu'il **ne** le sache. I it.have done without that.it NEG it know.sbJv 'I did it without knowing it.' $\Box \neg p$ The lexical items *°être exclu (que)* 'it is excluded (that)' and *°impossible (que)* 'impossible (that)' in French and *páichú* 'exclude'; *°bù-kěnéng* 'be impossible' in Mandarin can trigger EN, at least when they are negated. These triggers denote necessity modal operators and $\neg p$ is thus true in all (accessible) worlds. (24) is an example from French and (25) from Mandarin.

(24) Et il n'est pas impossible que Leopold n'ait lassé And it NEG-is NEG impossible that Leopold NEG-have.suBJV tired Grimm ...
Grimm ...
'And it is not impossible that Leopold tired Grim ...' (J. et B. Massin,

Mozart, Fayard, 1970, p. 241; cited in Muller 1991:375)

tǎolùn le rénlèi duì dìqiú de zhè běn shū yingxiang ... (25)This CLF book discuss PFV human to earth POSS influence wǒ juédé běn shū lǐ suǒ miáoshù de zhuàngkuàng think this book in pass describe POSS situation Τ bùkěnéng bú huì fāshēng, érqiě fāshēng búshì be.not.the.case impossible NEG will occur and occurrence jīlù huì suízhe shíjiān ér zēngjiā. time probability will with then increase 'This book discussed human influence on earth ... I think it is not impossible for the scenarios described in this book to arise; the possibility of their occurrence will increase with time.' (https://tinyurl. com/y26tmzqa, accessed on 2018-05-05)

 $\neg p$ in Suppositive Worlds This class of triggers involves conditional operators. The meaning of UNLESS entails that if $\neg p$ then q. $\neg p$ is thus true in worlds that are consistent with the hypothesis (what we call *suppositive worlds*).
Expletive Negation in English, French, and Mandarin

- (26) Nous n'interviendrons pas dans cette affaire à moins qu'on we NEG.interfere NEG in this affair to less that.INDEF
 ne nous le demande.
 NEG us it demand.SBJV
 'We won't interfere in this affair unless they ask us to do so.'
- qíshí chúle rúguð bú zhuàng-jìn jiàshǐcāng wēixié (27) jīzhǎng in.fact except if NEG hit-into threaten pilot cabin ānquán zhīwài, niǎo jī yìbān zàochéng fēijī jiégòuxìng de POSS safety outside bird hit normally cause plane structural sǔnshāng kěnéngxìng bú dà. de POSS damage probability NEG big 'In fact, unless a bird hits the cabin and threatens the pilot's safety, the probability that a bird's hit causes structural damage to a plane is normally not high.' (https://tinyurl.com/y6267pzh, accessed on 2018-05-05)

 $\neg A(Y, d')$ at w_0 Comparatives differ from all previous triggers in that the two propositions that are activated are not duals of each other (p and $\neg p$). The meaning of comparatives (X is Q-er) than (Y is Q) involves a comparison of degrees: there are degrees D such that X and Y are both Q to degree D and there are also (higher) degrees D' such that only X, but not Y, is Q to degree D'. Both [Q(X) to degree D'] and [not [Q(Y) to degree D']] are therefore true. Comparatives are the only EN-triggering context in which the two activated propositions of opposite polarity involve predications over distinct entities (in (28), the addressee and speaker, respectively) and where the two propositions of opposite polarity are true in the same world(s) or at the same time interval(s).

(28) Tu es plus grand que je **ne** suis. you are more tall than I NEG am 'You are taller than I am.'

Comparatives trigger EN in their complement clause in quite a few

languages aside from French. But Mandarin comparatives do not, as only NPs are allowed after the marker of comparison bi.

Table 1 summarizes the different classes of triggers and the negative proposition whose activation leads to the occurrence of EN, using one predicate as a representative of the class.

| Predicate | Negative proposition |
|-------------|---|
| FEAR | $\neg p$ in worlds consistent with X's desires |
| REGRET | $\neg p$ in worlds consistent with X's behavioral standards |
| DENY | $\neg p$ in worlds consistent with X's belief |
| FORGET | $\neg p$ in w_0 |
| BEFORE | $\neg p$ at reference time |
| IMPOSSIBLE | $\neg p$ in all accessible worlds |
| UNLESS | $\neg p$ in suppositive worlds |
| COMPARATIVE | $\neg A(Y, d')$ |

 Table 1 Classes of EN-Triggers and the Corresponding Negated Argument

 Proposition

6 Low-Entrenchment EN in English

Our model of EN predicts that EN occurs more often than typically recognized because grammarians and scholars overlook low-entrenchment EN. In fact, we predict that the same contexts should favor EN across languages. In this section, we test this prediction with data from English. English descriptive grammars (e.g., Quirk et al. 1985 or Huddleston and Pullum 2002) make no mention of EN, and prescriptive grammars do not allow sentences that include phrases like *refuse not to surrender* (meaning the same as *refuse to surrender*). However, some of Horn's (2010) examples suggest that (low-entrenchment) EN occurs in English. His examples include what would be EN-triggers in French and Mandarin such as *keep from, avoid, miss, hold back,* and *since*. There is, it seems, a parallel between the contexts that allow low-entrenchment EN in English and the contexts that allow the grammatically licensed negative marker *ne* in French. In this section, we provide examples that show that EN can occur in English in all of

the same contexts listed in table 1, thus confirming the prediction of our model that the causes of EN are of general applicability.

FEAR

(29) "It was a bit painful but the shock, once that wore out I really felt it. I was just worried that hopefully there wasn't a break in there and lucky enough there was no fracture, just a dislocation," Moa said. (https://tinyurl.com/y3y6d7ro, accessed on 2018-05-25)

REGRET

(30) I don't know why on earth he said that as I was the only person who stood by him when all others blamed him for not having a decent job. I always thought he was the one for me and at this point of time I really regret that I **shouldn't** have gone for him. I cried like anything for months and eventually got over the past only because of my few best friends. (https://tinyurl.com/y4f8a7gl, accessed on 2018-06-25)

DENY

(31) He also denied that he didn't mock a disabled reporter. there is proof of him doing this. He's such a liar. (https://tinyurl.com/ y4btqskr, accessed on 2018-05-04)

FORGET

(32) Then the worst happened. He forgot **not** to pick up after placing a pizza order, and there, on the other end of the line, was the whisky-voice of the sister, H, down in Palo Alto. (*Corpus of Contemporary American English*, example from Lois Taylor 'Are you there?' in *Southwest Review*, volume 92, 2007, pp. 82–88)

BEFORE

(33) When we realized we had time for one more song (backstage), we decided it had to be "Human Stain". Then at the end of the night we found out that we didn't play Karma and that felt so weird. It has been such a long time since we didn't play that song. (https://tinyurl.com/y4apwdm7, accessed on 2018-05-25)

IMPOSSIBLE

(34) It is not impossible that some aspect of sound-making efficiency might **not** have played into the mechanism of natural selection during the history of the species (E. Lenneberg, *Biological Foundations of Language*, 1967; cited in Horn 2010:127)

UNLESS

(35) He escorted the girls nearly everywhere they went, except when he was **not** busy with his own friends, who were quite fond of gambling and horses, like most young men-except perhaps Nathaniel Sheridan, who was too concerned with managing his father's many estates ever to stop for a game of whist or bagatelle. (https://tinyurl.com/ y47ywxy4, accessed on 2018-05-25)

COMPARATIVE

(36) This facilitated my work more than you will **never** know. (J. C. Smith, *Black Firsts: 4,000 Ground-Breaking and Pioneering Historical Events,* 2012)

Interestingly, we did not find an example where the expletive negative marker is simply *not* with either FEAR or REGRET. Additional words were present: in the case of FEAR, *hopefully* and in the case of REGRET, *should*. The use of these additional words echoes the use of specialized negations in Mandarin, the imperative negative marker *bié* in the case of FEAR and the modal negative marker *bù-gāi* 'shouldn't' in the case of REGRET. We surmise that the presence of words that evoke the modal contexts where $\neg p$ is true facilitates the occurrence of low-entrenchment EN in both cases.

7 Discussion

At the heart of the model of EN we have proposed in this paper is the hypothesis that the meaning of some words leads to a strong enough activation of $\neg p$ alongside its dual p that an EN marker surfaces. Because our model relies on language-general production mechanisms, it predicts EN to show up in roughly the same range of contexts across languages, provided near translation equivalents are considered and the syntax of those near translation equivalents does not ex-

clude (or strongly disfavors) the occurrence of EN. More specifically, since licensing of EN is conditioned on two language-independent factors, activation of inferred concepts in language production and the meaning of certain verbs, adverbs, and adpositions, our model predicts that the occurrence of EN should be universally possible for the eight classes of triggers listed in table 1.

Now, as we stressed in section 4, our prediction does not pertain to highly entrenched EN, but to all EN uses found in a language, low-entrenchment as well as highly entrenched EN. Only when lowentrenchment EN is included in the mix does the broad range of EN-triggers discussed in, say, Muller 1991 cease to seem an idiosyncrasy of French. We tested the cross-linguistically validity of our hypothesis by comparing the range of semantic contexts that license EN in French and Mandarin, on the one hand – two languages that are known for their relatively broad use of highly entrenched EN and English - a language that purportedly does not include highly entrenched EN uses. We found the range of triggers to be almost identical across all three languages. French, Mandarin, and English all include EN of some sort (highly entrenched or low-entrenchment EN) for all eight classes of triggers and exceptions can be explained by language-internal factors (e.g., the syntax of the relevant potential triggers). In a follow-up study (Jin & Koenig 2019), we corroborated this finding by a comprehensive survey of two more genealogically unrelated languages, Arabic and Zarma-Sorai. Almost all the same predicates can trigger some form of EN in these languages too. Again, exceptions are the result of language internal, mostly syntactic, factors. The uniformity of the range of triggers across the languages we looked at (65 are listed in Jin & Koenig 2019) further validates our methodology to include in the purview of our survey not only highly entrenched EN uses, but also low-entrenchment EN uses. English is then not the odd man out anymore and additional triggers, which would be expected from a semantic point of view, can be recognized in both French and Mandarin. As we mentioned, about half of the triggers we list and we found attested examples of were not listed as

EN-triggers in the previous Mandarin literature, most probably, we surmise, because grammarians tend to focus on highly entrenched EN uses.

We end this paper with a discussion of the fact that it is much easier to find EN with some triggers when the matrix clause is negated or questioned (indicated with ° in section 5). This observation has already been made in French grammars (e.g., Grevisse 1936) and studies of negation and EN (e.g., Muller 1991, Larrivée 2004). In fact, when we were investigating low-entrenchment EN in English, strings like 'not despair of not', 'not deny that he didn't', or 'not advise against not' yielded more examples of EN than the corresponding non-negated verbs. In some cases, the required presence of a negation in the matrix clause is somewhat puzzling semantically. Consider the need to negate the adjective impossible 'impossible' in French or the verb bù-kěnéng 'be impossible' in Mandarin. Since $\neg \Box \neg p \Leftrightarrow \diamond p$, it is odd that negating the predicate IMPOSSIBLE boosts the likelihood of occurrence of an EN, since, semantically, $\diamond p$ should activate $\neg p$ less than $\Box \neg p$. Our production model of EN provides, we believe, a possible explanation for this otherwise puzzling fact.

The critical property of both questioned (?p) and negated $(\neg p)$ propositions for our purposes is that they evoke p as well as $\neg p$ (see Hamblin 1973 for questions and Ducrot 1980 for negation). Neither Hamblin's approach to the meaning of questions (that it denotes the set of its answers) nor Ducrot's approach to the meaning of negation (that it evokes the assertion of p and $\neg p$ by two distinct enunciators) is uncontroversial. For our purposes it suffices that questions and negations have been argued to evoke (to use a minimally committing term) both the trigger and its negation. The concurrent activation of the trigger and its negation is, we suggest, what leads to a stronger activation of $\neg p$ and thus increases the likelihood of the occurrence of EN. For space considerations, we only discuss one case here.

Consider French *nier* and *impossible* where the negation is not just entailed by the meaning of the predicates, it is part of the mean-

ing of the predicates. So, the lexicalization of 'not' in $\neg p$ is already achieved by the lexical selection of *nier* and *impossible* and blocks the occurrence of a separate negative marker in the complement clause. But when DENY or IMPOSSIBLE are negated, we have a clash between the lexical meaning of each predicate (which lexicalizes the 'not' of $\neg p$) and the compositional meaning of *ne pas nier* or *n'être pas impossible* which is semantically incompatible with negating the argument proposition (again, $\neg \Box \neg p \Leftrightarrow \Diamond p$). The competition between the lexical meaning of the trigger that activates $\neg p$ and the compositional meaning, which does not, is what leads to EN. Our account of the occurrence of EN when *nier* and *impossible* are negated is similar to what Horn (2010) describes as *triplex negatio confundit*, that is, instances where a triple negation conveys a positive.

Clearly, our explanation of why some triggers strongly favor or require the presence of a matrix negation or question for EN to occur in the complement clause is quite speculative at this point. But it points to an interesting consequence of embedding our semantic account of the contexts where EN is found in a language production model: EN triggers can vary in strength, that is, in terms of the degree to which $\neg p$ is activated (e.g., because speakers are more or less likely to infer $\neg p$). When the activation of $\neg p$ is relatively weak, a boost might be needed for the activation of $\neg p$ to be strong enough to lead to the occurrence of an EN marker. Negating or questioning the trigger may be one such boost.

Acknowledgments We thank members of the Psycholinguistics Laboratory at the University at Buffalo, the audience of the Colloque de Syntaxe et Sémantique de Paris, Jürgen Bohnemeyer, Matthew Dryer, Karin Michelson, Robert Van Valin, and an anonymous reviewer for their invaluable comments. We would also like to extend our heartfelt thanks to Chris Piñón for his invaluable help.

References

Abels, K. (2005). "Expletive negation" in Russian: A conspiracy theory. *Journal of Slavic linguistics*, 13, 5–74.

- Cao, J. (2007). Xiànyú fŏudìng de yǔyòng rènzhī fēnxī. Master's thesis, Capital Normal University, Beijing.
- Damourette, J., & Pichon, E. (1911–1933). *Des mots à la pensée: Essai de grammaire de la langue française*. Paris: d'Artrey.
- Del Prete, F. (2008). A non-uniform semantic analysis of the Italian temporal connectives *prima* and *dopo*. *Natural Language Semantics*, *16*, 157–203.
- Dell, G. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, *93*, 283–321.
- Dindelegan, G. (2013). A Grammar of Romanian. Oxford University Press.
- Ducrot, O. (1980). Les mots du discours. Paris: Les Editions de Minuit.
- Espinal, M. T. (1992). Expletive negation and logical absorption. *The Linguistic Review*, *9*, 333–358.
- Espinal, M. T. (2000). Expletive negation, negative concord and feature checking. *Catalan Working Papers in Linguistics*, *8*, 47–69.
- Etxepare, R. (2003). Negation. In J. I. H. et al. (Ed.) *A Grammar of Basque*, (pp. 387–421). Berlin: Walter de Gruyter.
- Gaatone, D. (1971). *Etude descriptive du système de la négation en français contemporain*. Geneva: Librairie Droz.
- Grevisse, M. (1936). Le bon usage. Gembloux: Editions J. Duculot.
- Hamblin, C. L. (1973). Questions in Montague English. *Foundations of Language*, 10, 41–53.
- Harves, S. (2002). Genitive of negation and syntax of scope. In *Proceedings* of *ConSOLE IX*, (pp. 66–110). Leiden: Kluwer.
- Heim, I. (1992). Presupposition projection and the semantics of attitude verbs. *Journal of Semantics*, *9*, 183–221.
- Hintikka, J. (1969). Semantics for propositional attitudes. In J. W. Davis, J. Hockney, & W. K. Wilson (Eds.) *Philosophical Logic*, (pp. 21–45). Dordrecht: D. Reidel Publishing Company.
- Horn, L. (2010). Multiple negation in English and other languages. In L. Horn (Ed.) *The Expression of Negation*, (pp. 111–148). Berlin: Mouton de Gruyter.
- Hualde, J. I., & de Urbina, J. O. (2003). *A Grammar of Basque*. Berlin: Mouton de Gruyter.
- Huddleston, R., & Pullum, G. (2002). *The Cambridge Grammar of the English Language*. Cambridge University Press.
- Inkova, O. (2006). La négation explétive: Un regard d'ailleurs. Cahiers

Ferdinand de Saussure, 59, 107–129.

- Jiang, L. (2008). Gàiniàn diéjiā yǔ gòushì zhěnghé—kěndìng fǒudìng búduìchèng de jiěshì. *Zhōngguó yǔwén*, 327, 5–19.
- Jin, Y., & Koenig, J.-P. (2019). A cross-linguistic study of expletive negation. Manuscript.
- Kratzer, A. (1981). The notional category of modality. In H.-J. Eikmeyer, & H. Rieser (Eds.) *Words, Worlds, and Contexts*, (pp. 38–74). Berlin: Walter de Gruyter.
- Langacker, R. (1987). *Foundations of Cognitive Grammar*, vol. 1. Stanford, CA: Stanford University Press.
- Larrivée, P. (1996). Pas explétif. Revue romane, 31, 19-28.
- Larrivée, P. (2004). L'association négative: Depuis la syntaxe jusqu'à l'interprétation. Geneva: Librairie Droz.
- Lau, J. H., Clark, A., & Lappin, S. (2017). Grammaticality, acceptability, and probability: A probabilistic view of linguistic knowledge. *Cognitive Science*, *4*1, 1202–1241.
- Le Bidois, G., & Le Bidois, R. (1935–1938). *Syntaxe du français moderne*. Paris: A. Picard.
- Lü, S. (1980). Xiàndài hànyǔ bābǎi cí. Beijing: The Commercial Press.
- Makri, M.-M. (2013). *Expletive Negation beyond Romance*. Master's thesis, University of York, York.
- Martin, R. (1984). Pour une approache sémantico-logique du *ne* dit explétif. *Revue de linguistique romane, 48,* 99–121.
- Muller, C. (1991). La négation en français. Geneva: Librairie Droz.
- Nchare, A. L. (2012). *The Grammar of Shupamem*. Ph.D. thesis, New York University, New York.
- Portner, P., & Zanuttini, R. (2000). The force of negation in *wh*-exclamatives and interrogatives. In I. Martinez, L. Horn, & Y. Kato (Eds.) *Negation and Polarity: Syntactic and Semantic Perspectives*, (pp. 201–239). New York: John Benjamins Publishing Company.
- Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A Comprehensive Grammar of the English Language*. London: Longman.
- Rubinstein, A., Sichel, I., & Tsirkin-Sadan, A. (2015). Superfluous negation in Modern Hebrew and its origins. *Journal of Jewish Languages*, 3, 165– 182.
- Searle, J. (1969). Speech Acts. Cambridge University Press.

- Song, H. (2012). *Zébèi lèi xiànyú fŏudìng de duōjiǎodù fēnxī*. Master's thesis, Central China Normal University, Wuhan.
- van der Wouden, T. (1994). Polarity and 'illogical negation'. In M. Kazanawa, & C. J. Piñón (Eds.) *Dynamics, Polarity and Quantification*, (pp. 16–45). Stanford: CSLI Publications.
- Vázquez Molina, J. (2006). *La negación expletive en francés*. Ph.D. thesis, University of La Rioja, Logroño.
- von Prince, K. (2012). *A Grammar of Daakaka*. Ph.D. thesis, Humboldt-Universität zu Berlin, Berlin.
- Wang, J. (2008). *Xiàndài hànyǔ xiànyú fǒudìng jí qí géshì yánjiū*. Master's thesis, Yangzhou University, Yangzhou.
- Wang, Z. (2012). Xiàndài hànyǔ tèshū fǒudìng xiànxiàng rènzhī yánjiū. Ph.D. thesis, Shanghai Normal University, Shanghai.
- Yoon, S. (2011). Rhetorical comparatives: Polarity items, expletive negation, and subjunctive mood. *Journal of Pragmatics*, 22, 133–166.
- Yuan, Y. (2012). Dòngcí nèi yĭnxìng fǒudìng de yǔyì céngcì hé yìchū tiáojiàn. *Zhōngguó yǔwén*, 437, 99–113.
- Zhang, Y. (2004). Xiàndài hànyǔ fùcí tànsuǒ. Shanghai: Xuelin Press.
- Zhu, D. (1959). Shuō chà-yìdiǎn. Zhōngguó yǔwén, 87, 435-437.
- Zovko Dinković, I., & Ilc, G. (2017). Pleonastic negation from a crosslinguistic perspective. *Jezikoslovlje*, 18, 159–180.

A Finer-grained Typology of Perfective Operators

Fabienne Martin • Zsófia Gyarmathy

Abstract We argue that completion and maximality (defined through the notion of self-connectedness) both need to be taken into account as potential parameters in analyses of perfectives. Stative predicates in the French *passé composé* and *passé simple* require self-connected maximality and completion, while event (including activity) predicates require event completion only.

Keywords perfective · imperfective · maximality · self-connectedness · event completion vs. cessation · French · Hindi · Mandarin

```
F. Martin, Universität Stuttgart/Humboldt-Universität zu Berlin,
https://sites.google.com/view/fabienne-martin
Zs. Gyarmathy, GNW Kft, Budapest, https://sites.google.com/site/zsofiagyarmathy/
```

In Christopher Pinon (Ed.), Empirical Issues in Syntax and Semantics 12, 187–216. Paris: CSSP. http://www.cssp.cnrs.fr/eiss12/ © 2019 Fabienne Martin & Zsófia Gyarmathy

1 On Existential Closure, the Perfective Aspect and Divisive Reference

Event-semantic analyses commonly assume that a (non-quantified) sentence expresses existential quantification over events, typically realized by existential closure in the derivation (e.g., Kratzer 1996), often assumed to be introduced by tense or aspect, especially in neo-Reichenbachian accounts and formalizations of Klein 1994. A typical neo-Reichenbachian definition of the perfective operator (PFV) is as follows (Bohnemeyer 2014), where *P* is a variable for an eventuality predicate, t_T a variable for the topic time, and *g* the variable assignment function parameter with respect to a model *M*:

(1)
$$\llbracket \mathsf{PFV} \rrbracket^{M,g} = \lambda P \exists e[\tau(e) \subseteq t_T \land P(e)]$$

The imperfective operator (IMPF) is assumed to express the inverse

relation between the topic time and the event time:

(2) $\llbracket \mathsf{IMPF} \rrbracket^{M,g} = \lambda P \exists e[t_T \subseteq \tau(e) \land P(e)]$ (Disregarding the imperfective paradox)

Another commonly assumed tenet in event semantics is that states are *divisive*, that is, they apply to all the parts of any eventuality they apply to. So if the event predicate P is divisive, then $\forall e, e'(P(e) \land e' \sqsubset e \rightarrow P(e'))$ holds, where \sqsubset is the strict mereological part-of relation over eventualities. An often dismissed consequence of the divisive reference of states and (1) is the following: whenever $t_T \subseteq \tau(e')$ for some e' in the denotation of a stative predicate P, it follows (because of divisive reference) that $\exists e[\tau(e) \subseteq t_T \land P(e)]$ holds. That is, IMPF(P) entails PFV(P) for stative Ps. Therefore, whenever an imperfective is true, a corresponding perfective sentence should be truthfully assertable.

However, this is not what we always observe. Suppose that Kim has been continuously sick since yesterday. In this scenario, sentence (3) (in which a clause with the present imperfective is followed by one with the simple past) is true. But the corresponding French example (4), whose second clause contains a perfective verb in the *passé composé* (PC), is not. This is clearly unexpected: the first clause states that $\exists e [\mathsf{now} \subseteq \tau(e) \land \mathsf{kim-be-sick}(e)]$ holds, and (by assumption), that yesterday $\subseteq \tau(e)$ holds (since *e* began yesterday). Given the divisiveness of states, it follows that $\exists e'[\tau(e') \subseteq \mathsf{yesterday} \land P(e')]$ holds (where *e'* is a part of *e*), and thus, a perfective sentence such as the second clause of (4) should be true if it expresses this meaning. But in fact, (4) is only acceptable if Kim's past state of sickness *e'* ceased in the past, followed by a present state of sickness *e*, and so that $\tau(e')$ and $\tau(e)$ are separated by a time interval *t* during which Kim is not sick. This suggests that the semantics in (1) and/or (2) is not correct.

- (3) Kim **is** sick, and she **was** already sick yesterday.
- (4) Kim **est** malade, et elle **a** déjà **été** malade Kim is.3sG sick and she be.PC.3sG already sick

hier. yesterday. 'Kim is sick, and she was already sick yesterday.'

The same problem can be approached from a different angle. Since for stative *P*s, IMPF(P) entails PFV(P), there is no entailment from PFV(P) to IMPF(P) or its negation. So if PFV(P) is true, IMPF(P) may be true or false as far as the definitions in (1) and (2) are concerned. But again, this is not what we observe. Data suggest that for stative *Ps*, if PFV(P) is true, then IMPF(P) is *false*, indicating a problem with (1) and/or (2). For instance, in contrast to (5) (in which a clause with the simple past is followed by one with the present imperfective), the corresponding French (6) from Schaden 2015, whose first clause contains a perfective verb in the PC and the second clause contains a present imperfective verb, is anomalous.¹ The same problem arises with a perfective verb in the *passé simple* (PS), as in (7).

- (5) There **was** a bar at the corner, and it still **is** there.
- (6) Il y **a eu** un bar au coin, #et il y **est** toujours. There be.PC.3SG a bar at-the corner and there is still. Intended: 'There was a bar at the corner, and it still is there.'
- (7) En 2000, Ana habita à Paris. #Et elle y habite in 2000 Ana live.ps.3sG in Paris and she there liveps.3sG toujours. still Intended: 'In 2000, Ana lived in Paris, and she still lives there.'

On a neo-Reichenbachian account like Schaden 2015, the contrast between (5) and (6) (and between (7) and its English translation) might be approached as follows. Since PFV requires the event time to be included in the topic time, asserting that the state continues to occur at utterance time leads to a contradiction. In contrast, the

¹See also Smith (1991:p. 195) for a similar observation on French. Note that an ingressive/inchoative interpretation in (6) and (7) is to our ears not possible.

English simple past is not a pure perfective and also has imperfective readings (cf. Comrie 1976), hence the absence of contradiction.

However, since for stative predicates, IMPF(P) may be true while PFV(P) is true, no contradiction should arise in (6) in a scenario where a present 'be-at-the-corner' state was already holding in the past. The same reasoning applies to (7). This problem cannot be circumvented by an appeal to Gricean implicatures and saying that the use of the perfective implies that the corresponding weaker imperfective is false, because that does not explain the *contradiction* in (6) observed by Schaden, as the problem does not vanish in presence of *en fait* 'in fact', which usually helps to cancel Gricean implicatures:

(8) Il y a eu un restaurant chinois dans ce there be.PC.3SG a restaurant Chinese in this quartier, #et en fait il est toujours là. neighborhood, and in fact it is still there Intended: 'There was a Chinese restaurant in this neighborhood, and in fact it is still there.'

Thus, existential quantification over events together with divisive reference leads to unacceptable results.

But could we not exploit the weakness of existential quantification and say that although PFV cannot be used for the French PC/PS (which don't have an imperfective use), it might be ideal for the English simple past, which is known to behave like a perfective with telic predicates (e.g., *John walked to the bank, #and he is still walking there*), and like an imperfective with atelic ones (recall (3) and (5))? More concretely, if we assume that the English simple past systematically satisfies the definition of PFV, we seem to predict exactly what we observe, namely that, when it comes to states, (i) IMPF(*P*) entails PFV(*P*) and (ii) if PFV(*P*) is true, then IMPF(*P*) may very well be true, too. However, this leads to another problem illustrated by the two potential readings of (9):

- a. topic time (*when*-clause) \subseteq event time (*be sick*) (Most salient reading)
- event time (*be sick*) ⊆ topic time (*when*-clause) (Marginal reading)

If we assume that the simple past in the main clause satisfies the definition of PFV (resulting in reading (9b)), we cannot account for the saliency of (9a). And notice that (9a) must be translated to French using the *imparfait* (which is an imperfective), while (9b) is preferably translated using the perfective PC. This strongly suggests that the English simple past is genuinely ambiguous, having both imperfective and perfective meanings.² However, even on an ambiguity analysis, if on its perfective reading the simple past encodes PFV, we arrive back at the question of why (9b) is only very marginally available, given that the perfective reading is also entailed by the (preferred) imperfective one.

So an analysis of PFV along the lines of (1) leads to problems both in the case of "pure" perfectives and the English simple past. The reason this issue has gone virtually unnoticed in the literature is that existential closure is actually mostly *disregarded* when talking about events and event times in the neo-Reichenbachian tradition and event semantic approaches in general. Authors, including Schaden (2015), often simply use the definite description "the eventuality". Thus, it is tacitly assumed that a sentence is about a specific eventuality, even though the formal analysis fails to do justice to this intuition.

²Of course, the translation facts alone do not help us select between an underspecification and an ambiguity analysis of the simple past, but together with the saliency of (9a), they are at least suggestive.

Still another view is that the English simple past is a pure past tense, and not an aspectual operator. For instance, de Swart (1998) claims that the English simple past is aspectually transparent in that it lets the lexical aspect "shine through" at the sentence level.

2 A Referential Approach to Aspect

Partee (1973) famously argued for a referential – as opposed to an existential quantificational – approach to tense. Based on cases of event anaphora, Grønn & von Stechow (2016) argued for an extension of Partee's idea to aspects, maintaining that both tenses and aspects have both a quantificational (indefinite) and a referential (definite) use. If, instead of existential closure over events, we use a variable whose reference is determined by the assignment function *g* (like that of t_T), as in (10), it becomes possible to capture the intuition that sentences are about specific events, and we can refer to *the* event.

(10) a.
$$\llbracket \mathsf{PFV} \rrbracket^{M,g} = \lambda P[\tau(e) \subseteq t_T \land P(e)]$$

b. $\llbracket \mathsf{IMPF} \rrbracket^{M,g} = \lambda P[t_T \subseteq \tau(e) \land P(e)]$

Given the lack of existential quantification, IMPF(P) will also no longer entail PFV(P) even in the case of predicates with divisive reference.

It is still not clear, however, why the perfective in (6) cannot refer to a (specific) past be-at-the-corner state which happens to be a proper part of a larger be-at-the-corner state that still holds. One could try to solve this problem by stating that in the domain of eventualities, the uniqueness requirement attached to definites translates into a constraint that forces the specific eventuality to be *maximal* (Filip 1999, Koenig & Muansuwan 2000, Altshuler 2014), that is, to be an eventuality that *ceased*. For if the state reported in (6) is the *unique greatest* past be-at-the-corner state in the context, it cannot be part of a larger (and therefore distinct) be-at-the-corner state. One of the problems with such an argument, however, is that it predicts all definite aspects to encode maximality, including the imperfective. However, the imperfective version of (6) is, of course, unproblematic.³

³Positing that the imperfective only has the indefinite reading in definition (2) is undesirable, since it can be anaphoric to a familiar event in languages like Russian (Grønn & von Stechow 2016) or French. This is, for instance, the case in the following example: *Hier, Pierre a dansé comme un fou. Il dansait avec un parapluie*.

3 Combining Event Maximality with Event Completion

We propose to employ maximality not as a requirement of definite aspect but that of PFV. Our claim is that the data in section 1 offer support to Altshuler and Filip's (2014) proposal that perfectivity amounts to a *maximality requirement* that is satisfied when an eventuality is a complete VP-eventuality⁴ or ceases to develop further towards a VP-eventuality in the actual world. Altshuler (2014) argued that the Hindi perfective encodes event maximality (but not event completion). The maximality requirement may be fulfilled in (11) (since the context leaves open the possibility that the eatingthe-cookie event ceased in the past), where the telic VP has a nonculminating, partitive use, but not in (12), since the second clause indicates that the eating-a-cookie event is still ongoing at utterance time. Altshuler (2014) proposes that this is at the source of the infelicity observed in (12).

- (11) maĩ-ne aaj apnaa kek khaa-yaa, (aur baakii I-ERG today mine cake eat-PFV and remaining kal khaũũgaa). (Hindi) tomorrow eat.FUT 'I ate my cake today, (and I will eat the remaining part tomorrow).' (Singh 1991)
 (12) maayaa-ne biskuT-ko khaa-yaa #aur use ab tak khaa Maya-ERC cookie-ACC eat-PEV and it now still eat
- Maya-ERG cookie-ACC eat-PFV and it now still eat rahii hai. (Hindi) PROG be.PRS Intended: 'Maya ate a cookie, and is still eating it.' (Altshuler 2014)

Note that Altshuler (2014) and Altshuler & Filip (2014) adopt a

^{&#}x27;Yesterday Pierre dance.PC.3SG like mad. He dance.IMP.3SG with an umbrella.'

⁴While it may not be the best terminology when it comes to atelics, we use "complete" as in Zucchi 1999 to express that the event falls under the respective predicate: e is complete wrt P iff P(e). In contrast to accomplishments, atelics (states and activities) are "complete" as soon as they begin.

maximality requirement that inherently refers to *stages* in the sense of Landman (1992), which incorrectly precludes all perfective operators from applying to states, which do not have stages. The data in section 1 warrant a relaxation of the maximality requirement from stages to parts. For the present purposes, the definition of maximality in (13) suffices. But in section 4.3, we will show that (13) is too strong, and offer a revised version of maximality.

(13) MAX(e, P) iff

(To be revised)

- a. *e* is a part of a possible *P*-event and
- b. it is not a proper part of any actual event that is part of a possible *P*-event.

With maximality included in the definition of PFV (as in the definitions in table 1), we can account for (6) and (7). We can also explain the marginality of (9b): IMPF(P) (as analyzed in (2)) no longer entails PFV(P) for predicates with divisive reference. The contrasts in (5)–(7) are due to the fact that the English simple past has an imperfective reading, while the French PC/PS do not.⁵

Importantly, we claim that, as in Altshuler & Filip 2014 for the Russian perfective, the maximality requirement of the English simple past and the French PC/PS does not *replace* the completion requirement they are traditionally associated with, but has to be *combined* with it to prevent the perfective aspect itself from leading to partitive readings of telics (recall *John walked to the bank, #and he still is walk-ing there.*) The cross-linguistic typology of Altshuler 2014 can thus be extended as in table 1: while the Hindi perfective encodes maximality only, the English simple past and the French PS/PC – just like the Russian perfective in Altshuler & Filip 2014 – encode maximality *and* completion. We call perfectives of the former type *weak perfec-*

⁵The PC has an imperfective reading when used as a universal perfect. But this use tends to require an adverbial like *toujours* 'always' or *depuis* 'since' in the PC sentence; see, for example, Schaden's (2007) example (102) *Depuis le début de l'hiver, Marie a été malade tout le temps* 'Since the beginning of the winter, Marie has been sick all the time'. No such adverbial is present in (6) (or (4)).

A Finer-grained Typology of Perfective Operators

| Perfective operator | Requires completion? | Requires maximality? | Semantics |
|---------------------|----------------------|-------------------------|-----------------------------------|
| Weak (Hindi, | No | Yes | [[PFV _M]] |
| Mandarin) | | | |
| Strong (French, | Yes | Yes | $\llbracket PFV_{C+M} \rrbracket$ |
| English, Russian) | | | |

Table 1 A typology of perfective operators (To be revised) $\llbracket \mathsf{PFV}_M \rrbracket = \lambda P \exists e[\tau(e) \subseteq t_T \land \mathsf{MAX}(e, P)], \text{ while}$ $\llbracket \mathsf{PFV}_{C+M} \rrbracket = \lambda P \exists e[\tau(e) \subseteq t_T \land \mathsf{MAX}(e, P) \land P(e)]$

tives, and those of the latter type *strong perfectives*. The typology will be revised in the last section.⁶

4 Three Problems Raised by Perfectives and Their Solution

4.1 Weak Perfectives and Complex Verbs

An issue raised by weak perfectives as defined in table 1 appears in languages like Hindi and Mandarin, where perfective accomplishments may have partitive readings with *simple verbs* (SVs) only, as in (11), or (14)–(17).

| (14) | John-ne fasal kaaT-ii . (Hindi) | | | | |
|------|--|--|--|--|--|
| | John-erg crop cut-pfv.sg | | | | |
| | 'John cut the crop (partly/entirely).' | | | | |
| (15) | John-ne draing miTaa-yii . (Hindi) | | | | |
| | John-ERG drawing erase-PFV.SG | | | | |
| | 'John erased the drawing (partly/entirely).' | | | | |

⁶Note that Altshuler (2014) and Altshuler & Filip (2014) adopt a purely event mereological approach on which viewpoint aspect operators are functions from eventuality predicates to eventuality predicates, while we here remain within the neo-Reichenbachian/Kleinian tradition introduced in section 1. However, nothing hinges on this choice with respect to the questions and analyses discussed here (e.g., the issues in section 1 arise for purely event mereological approaches, too).

(16) Wǒ zuótiān xiě le gěi Zhāngsān de xìn, kěshì I yesterday write PFV to Zhangsan DE letter but méi xiě wán. (Mandarin) NEG.PFV write finish 'Yesterday, I wrote a letter to Zhangsan, but I didn't finish (writing it).' (Koenig & Muansuwan 2000)
(17) Yuēhàn shāo le tā-de shū dàn gānhěn méi

(17) Yuēhàn shāo le tā-de shū, dàn gēnběn méi Yuehan burn PFV 3SG-DE book but at.all NEG.PFV shāo-zháo. (Mandarin) burn-ignited 'Yuehan burned his book, but it didn't get burnt at all.' (Martin et al. 2018)

These readings are unavailable when applied to *complex* or *compound verbs* (henceforth CVs): compare, for example, (11) (which contains a SV) with (18) (which contains a CV). The contrast between (17) and (19) illustrates the same point.

(18) maĩ-ne kek khaa liya, #jo bacaa hai wo raam I-ERG cake eat take.PFV what remain is that Ram khaayegaa. (Hindi) eat.FUT Intended: 'I ate the cake (completely), and Ram will eat the rest.' (Singh 1994)

 (19) Yuēhàn shāo-zháo le tā-de shū, #dàn gēnběn méi Yuehan burn-ignited PFV 3SG-DE book but at.all NEG.PFV shāo-zháo. (Mandarin) burn-ignited 'Yuehan burned his book, but it didn't get burnt at all.' (Martin et al. 2018)

CVs are formally composed of two roots that can often both be used as independent main predicates outside the CV, where V1 is either a verbal root describing an event, and V2 describes the result state or the right boundary (*telos*) of a V1-event, or is a light verb that has lost independent meaning. Both in Hindi and Mandarin, perfective CVs entail that the V1-event either has reached its right boundary (was completed) and/or has triggered a V2-result; see Singh 1991, 1994 and Altshuler 2014 for Hindi; Li & Thompson 1981, Lin 2004 and Chief 2008 for Mandarin.⁷ If, as we saw in section 3, the perfective is weak in these languages, this is at first sight unexpected, given that the aspectual operator should be able to extract an initial proper part of an eventuality satisfying the CV.

Focusing on Hindi, Altshuler (2014:746) solves the issue by assuming that Hindi has two different perfective operators (hence his labels '^{SV}PFV' vs. '^{CV}PFV'). He assumes that SVs combine with the weak perfective analysed as ^{SV}PFV (which encodes event maximality only), while CVs combine with the strong perfective analysed as CV PFV (which encodes event completion). If possible, one might prefer to avoid this solution, however, given that the same morphology (namely, -(y)aa/ii) is used to express perfectivity with both types of predicates. The same problem arises in Mandarin, where perfectivity is encoded by one and the same verbal morpheme *le*.

We propose that Piñón's (2011) account for the semantic differences between English simple vs. particle verbs such as *eat* vs. *eat up* can be fruitfully extended to this semantic contrast between perfective SVs and CVs in Mandarin or Hindi. Piñón's point of departure is a suggestion made by Higginbotham (2000) that verbs may either denote predicates of events, or of *ordered pairs of events*. Piñón proposes that verbs like *eat* may come in two variants, as in (20a) and (20b), where *b* is a variable for boundary events, and 'V+' indicates that *b* is the boundary of *e*, see (20c).⁸

⁷But see Koenig & Muansuwan 2000 on the view that Thai perfective CVs only strongly imply rather than entail event completion.

⁸Note that we left out from Piñón's (2011) representations the internal argument and the thematic role he assumes to be introduced by the verb. Instead, we assume (in the spirit of Distributed Morphology) that the verbal root *eat* – and its Mandarin or Hindi counterparts – introduces an event argument only, while the internal argument and the patient thematic role are introduced by a separate head.

(20) a. eat $\rightsquigarrow \lambda e.eat(e)$ b. eat up $\rightsquigarrow \lambda \langle e, b \rangle.eat^+(\langle e, b \rangle)$ c. $\forall \langle e, b \rangle (\mathbf{V}^+(\langle e, b \rangle) \rightarrow b = right-boundary-of(e) \land b \sqsubset e)$

The strategy we propose consists in treating Hindi and Mandarin CVs as denoting predicates of ordered pairs of events, like eat up does in (20b), while SVs denote predicates of events. If V2 is an achievement verb, Piñón's analysis hardly needs any modification in order to be extended to CVs in these languages. Interestingly, many so-called "vector verbs," that is, predicates or auxiliaries used in the V2-position of CVs in Indo-Aryan (Hindi, Marathi), Dravidian (Tamil, Malayalam), and Turkic (Tatar) languages have a literal meaning expressed in English by achievement(-like) verbs such as finish, leave, go, come, reach, put, throw, fall - see Maisak 1999 for a rich inventory of these verbs. This corroborates the idea that V2 is a boundary-denoting (achievement) predicate in at least some CVs. For instance, Hindi khânâ 'eat' can be attributed the same meaning as eat in (20a), and the corresponding CV (formed with the light verb li 'take') the meaning in (20b). Similarly, the Mandarin SV guān 'close' can be analysed as in (21a), and the corresponding CV guānshàng 'close up', containing the movement verb shàng 'rise', as in (21b), where ' $\mathbf{V}^{-}(b, s)$ ' (in the spirit of (20c)) indicates that *b* is the left boundary of *s*, and, by the axiom about events and their caused result states in (21c), we also assume the right boundary of *e* in (21b).

- (21) a. $gu\bar{a}n$ 'close' $\rightsquigarrow \lambda e. \exists s(\mathbf{cause}(e, s) \land \mathbf{closed}(s))$
 - b. $gu\bar{a}n$ -shàng 'close up' $\rightsquigarrow \lambda \langle e, b \rangle$. $\exists s(cause(e,s) \land closed^{-}(b,s))$
 - c. $\forall \langle e, b \rangle \forall s((cause(e, s) \land \mathbf{V}^{-}(b, s)) \rightarrow (\mathbf{V}(s) \land b = right-boundary-of(e) \land b = left-boundary-of(s)))$

When V₂ is a state predicate, or a non-punctual event predicate, Piñón's analysis has to be modified further. A different relation (typically a causal relation in so-called resultative verbal compounds) holds between the eventualities respectively expressed by the first and second members of the ordered pair. For instance, the simple and complex Mandarin variants of the causative verb *burn* can be analysed as follows.⁹

(22) a.
$$sh\bar{a}o$$
 'burn' $\rightsquigarrow \lambda e.\exists s(cause(e,s) \land burnt(s))$
b. $sh\bar{a}o$ - $zh\dot{a}o$ 'burn-ignited' $\rightsquigarrow \lambda \langle e, s' \rangle.\exists s(cause(e,s) \land burnt(s) \land cause(e,s') \land ignited(s'))$

The price of such an analysis is that the denotations are more complex type-logically, but the contrast in terms of completion entailment can now be accounted for. When a weak perfective applies to a CV denoting a predicate of ordered pairs of events, it applies to an ordered pair of events that is maximal with regard to the CV, that is, MAX($\langle e, e' \rangle$, V1.V2). For instance, the perfective form of *shāo-zháo* 'burn-ignited' receives the semantics in (23) (and the MAX operator must now, of course, be extended to accept event pairs as argument):

(23) $\mathsf{PFV}[sh\bar{a}o\text{-}zh\acute{a}o] \rightsquigarrow \exists \langle e'', s'' \rangle (\tau(\langle e'', s'' \rangle) \subseteq t_T \land \\ \mathsf{MAX}(\langle e'', s'' \rangle, \lambda \langle e, s' \rangle. \exists s(\mathsf{cause}(e, s) \land \mathsf{burnt}(s) \land \\ \mathsf{cause}(e, s') \land \mathsf{ignited}(s'))))$

The maximality requirement is satisfied either by a pair of events which is a complete V1.V2-pair of events, or by a pair of events which is an incomplete V1.V2-pair of events that ceases to develop further in the actual world. But crucially, an incomplete *pair* of events must still be a pair of (incomplete) events; an ordered pair of events cannot be made up of an initial part of the first member of the pair only. Therefore, a perfective CV requires at least a part of a V2-boundary or of a V2-result to occur. Since a boundary is an atomic object without proper parts (Piñón 1997), CVs whose V2 is boundary-denoting entail event completion, which is the result desired. When V2 is a

⁹We follow Martin et al.'s (2018) proposal according to which Mandarin SVs such as *shāo* 'burn' are causative (bi-eventive) predicates. Also, note that predicates are often categorically underspecified and can be both used as a verb or an adjective; we assume that in a causative CV, V1 is used as a verb, introducing an event leading to a state *s*, and V2 introduces a state *s*'.

state predicate, the ensuing perfective CV minimally entails a proper part of a V2-state, which is again what we observe. In both cases, event completion is ensured with CVs despite weak perfectivity, due to the requirement that a V2-event fragment be instantiated.¹⁰ By contrast, the perfective of a causative SV does not entail the existence of a state *s* satisfying the property encoded by the SV, see (24).

(24) $\mathsf{PFV}[sh\bar{a}o] \rightsquigarrow \exists e'(\tau(e') \subseteq t_T \land \mathsf{MAX}(e', \lambda e. \exists s(\mathbf{cause}(e, s) \land \mathbf{burnt}(s))))$

4.2 Ongoing Readings after Perfective Activities

As we saw in section 1, states in the perfective do not allow for an ongoing eventuality in French. We have also noted that the addition of *en fait* 'in fact' does not lead to acceptability, indicating that cessation is not simply an implicature of the perfective. The main data can be illustrated as follows (note that cessation is entailed both with predicates denoting states that normally last a long time, as in (25a)-(25c), and predicates denoting states whose lifespan can be very short, as in (25d)):¹¹

 (25) a. L'année passée, Pierre a habité à Paris #et il y the.year last Pierre live.PC.3SG in Paris and he there habite toujours. live.PRES.3SG still

 Lulu shāo-zháo le nèi-běn shū, dàn shū méi wánquán zháo.
 Lulu burn-ignite PFV that-CL book but book NEG.PFV completely ignite 'Lulu burned that book, but the book didn't get completely burnt.'

¹¹However, as already observed by Smith (1991) about a similar French example, (25d) sounds better if we admit the existence of an iteration of sick-states, which we will account for below.

¹⁰When the Mandarin CV contains a gradable result predicate V₂, it is left open whether or not the result satisfies V₂ to a maximal degree. What is required is that an eventuality of the V₂-type occurs, exactly as expected if MAX combines with CVs having a semantics such as (22b): (22b) leaves open whether the result state *s'* satisfies the predicate *ignited* to a maximal degree or not. For instance, Martin et al. (2018) report the following example to be non-contradictory:

Intended: 'Last year, Pierre lived in Paris, and he still lives there.'

- b. Il y a eu un restaurant chinois dans ce There be.PC.3SG a restaurant Chinese in this quartier, #et en fait il est toujours là. neighborhood, and in fact it be.PRES.3SG still there Intended: 'There was a Chinese restaurant in this neighborhood, and in fact it is still there.'
- c. Marie a été plus grande que son frère (#)et Marie be.PC.3SG more tall than her brother and elle l'est toujours. she that.be.PRES.3SG still Intended: 'Marie was taller than her brother, and still is.'
- d. Ce matin Pierre a été malade (#)et il this morning Pierre be.PC.3SG sick and he l'est toujours. that.be.PRES.3SG still Intended: 'Pierre was sick this morning, and he still is.'

We have argued that these data, given the divisive reference of stative predicates, present a challenge to most traditional approaches that assume that aspect introduces existential quantification over events and PFV simply requires event completion. However, the question may arise as to what the case is with activities, which are also assumed to have divisive reference, albeit only down to minimal parts. The surprising observation is that in the case of activities in the PC-perfective, an ongoing-event interpretation *is* possible, as the following felicitous French examples show:

 (26) a. Ce matin Pierre a travaillé et il travaille this morning Pierre work.PC.3SG and he work.PRES.3SG toujours. still
 'This morning, Pierre was working (lit.: worked) and he is still working.'

- b. Ce matin Pierre a travaillé et il this morning Pierre work.PC.3SG and he n'a pas arrêté depuis lors.
 NEG.stop.PC.3SG since then 'This morning, Pierre was working (lit.: worked) and he hasn't stopped working since then.'
- c. Maya a mangé des cookies (ce midi), et elle en Maya eat.PC.3SG of cookie.PL (this noon) and she of mange toujours (depuis lors).
 eat.PRES.3SG still (since then)
 'Maya was eating (lit.: ate) cookies (at noon today), and she is still eating (since then).'

Crucially, an ongoing-event interpretation in these cases is possible both with and without the interruption of the event that makes the perfective true. (And note that in English, this ongoing-event interpretation raises problems with the simple past, as already noted in Smith 1999, which justifies our translation with a progressive). Moreover, this contrast between states and activities holds up even if only a *minimal part* of the relevant activity has been completed, as evidenced by the following pair of examples (and note that (27b) is fine once Armstrong is back in his spaceship).

- (27) (Context: Neil Armstrong took his first step on the moon and is still walking.)
 - a. Armstrong **a marché** sur la lune! Armstrong walk.PC.3SG on the moon 'Armstrong has walked on the moon!'
 - b. #Armstrong **a été** sur la lune! Armstrong be.PC.3SG on the moon Intended: 'Armstrong has been on the moon!'

This contrast between states and activities, and the activity data, in particular, are highly puzzling in view of our proposed analysis of the French PC as involving both completion and maximality: if, for

instance, in (27a), the PC form for 'walk on the moon' can only be made true by maximal walkings-on-the-moon, how come it can be felicitously used in a situation when it is obviously made true by a non-maximal activity? Should we give up, given this data, our proposal that PC requires maximality? While such a move would certainly explain the felicity of (27a) and its ilk, we would then be hard put to account for the data concerning states and the issues detailed in section 1. In particular, if the PC did not require maximality, how can the infelicity of (25) and (27b) be explained? The problem is not specific to French, since in Mandarin and Hindi, too, an ongoingevent context is less problematic for perfective activities than perfective states. For instance, (28) in Hindi is far from being categorically rejected by our informants¹² (although it is still marked for most of them), and the same is true of sentences like (29) in Mandarin.¹³

(28) mayaa-ne khanaa **kha-yaa**, (?)aur ab tak khaa rahii Maya-ERG food eat-PFV, and now still eat PROG

¹²*Pace* Singh (1991:11), who categorically rejects ongoing readings with perfective activities (see her (25)).

¹³To our knowledge, the case of perfective statives with pure stative readings is not very much discussed in Hindi and Mandarin. In Mandarin, the combination of the verbal *-le* with stative predicates very often forces an inchoative interpretation: see, for example, Lin 2004 (and note that although this inchoative reading is very often *optional* in French, the literature tends to focus exclusively on it to the detriment of the pure stative reading). In the rare cases where perfective states have a truly stative meaning in Mandarin, the ongoing interpretation is odd, see, for example, (i). As for Hindi, our informants converge on the view that a perfective stative sentence such as (ii) is very marked, too.

- (i) Shàng gè yuè, Lùlu zài Bālí dāi le shí-tiān, #tā hái dāi zài Bālí. last CL month Lulu at Paris stay PFV ten-day 3SG still stay at Paris Intended: 'Last month, Lulu stayed in Paris for ten days, and she still is staying in Paris.'
- (ii) mayaa is hotel me ruk-ii, #aur ab tak ruki hai.
 maya this hotel in stay.PFV.SG and now still stay be.PRS
 'Maya stayed in this hotel, and she is still staying there.'

hai. be-prs 'Maya ate, and she is still eating.'

(29) Lùlu jīntiān zǎoshàng pǎo le bù, (?)yìzhí pǎo Lulu today morning run PFv step all.along run dào xiànzài.
up.to now 'Lulu ran this morning, and she has been running until now.'

We therefore propose to retain our previous account of the French PC (as well as Mandarin *le* and Hindi -(y)aa/ii) as requiring maximality, and instead claim that some difference between states and activities should account for the difference in the availability of an ongoing-event interpretation in their cases (although we ultimately prefer the alternative solution provided in the next section).

We here tentatively propose that the crucial factor in this respect is the not entirely homogeneous nature of activities, as opposed to states. Since at least Taylor 1977 and Dowty 1979, the received view is that states have divisive reference down to instantaneous parts, while activities do so only down to some small, "minimal" parts that are sufficiently big enough to allow them to be classified under the relevant predicate. For instance, walk can only hold true of at least one step-sized events, while states like be hold at all subintervals (down to instants) of any interval at which they hold. We hypothesize that French speakers are sensitive to this difference to the extent that they are able to construe the end of a minimal part of an activity and the subsequent tiny pause until the next substantial minimal part commences as cessation. More concretely, the idea is that in an example such as (26a), the activity *e* reported in the first clause may be conceived as maximal with respect to the predicate travailler 'work': e does cease in the past and is therefore not a proper part of a bigger and still on-going working-activity, but is rather followed by another activity e' of the same type.

In fact, Rothstein (2004) already suggested a systematic ambiguity

of referring to minimal parts only, or to members of the set formed from them through closure under join. Albeit she argued (for English) that only "naturally atomic" activities can access these minimal parts, while "ordinary" activities unambiguously denote the set (let us call it P following Rothstein) formed through closure under join from them, there appears to be no reason to exclude native speakers from accessing the minimal parts. In fact, P of course includes these minimal parts (by definition, since they are the smallest events that can make true an activity of a given type). As such, we expect that predicates like marcher 'walk' can refer to minimal events like steps, and given this, there is no reason to exclude native speakers from construing tiny pauses following them as intervals at which the activity ceases to hold - which satisfies the requirement of maximality. In contrast, since states are true down to instants, no such pause is encountered which would constitute enough grounds to assume the cessation of the state at any moment in time (up until the state does cease).14

Although such an explanation may prove difficult to validate (and we will see in the next section that maximality as defined in (13) raises an even more serious problem with cumulative predicates), one piece of evidence that we take to at least weakly reinforce it concerns verb iterations. Verb iterations in French can be used to signal unexpectedly long eventualities, and are readily available for activities, as the following example shows:

 (30) Pierre travailla, travailla, travailla dans son Pierre work.PS.3SG work.PS.3SG work.PS.3SG in his bureau. office 'Peter was working, and working, and working in his office.'

¹⁴And note that in a context such that a pause may be conceived between different **P**-states, the ongoing reading sounds much better. For instance, (25d) is acceptable if one assumes the occurrence of two different states of sickness (although the adverbial *de nouveau* 'again' fits better this context than *toujours* 'still').

In contrast, this construction is not acceptable in the case of states:¹⁵

(31) #Pierre resta, resta, resta dans son bureau. Pierre stay.Ps.3SG stay.Ps.3SG in his office Intended: 'Peter was staying for a long time in his office.'

One potential explanation for this difference between activities and states with respect to verbal iteration may be based on the sensitivity of French speakers to tiny pauses between minimal parts of activities. Thereby, a longer activity (of, say, working) may be divided into some smaller chunks, and the numerousness of these chunks indicated through verbal iteration (note that it is not uncommon in languages to use iteration to signal plurality or intensification).

A second piece of evidence that might be taken to support our proposal that French speakers can construe minimal activities as maximal activities comes from counting facts. Let as assume that we can felicitously utter the following iterated form, suggesting a long drawing session:

(32) Sascha a dessiné (et) dessiné (et) dessiné. Sascha have.3sG draw.PC (and) draw.PC (and) draw.PC 'Sascha was drawing, and drawing, and drawing.'

In this context, *plusieurs fois* 'several times' is a true answer to the following question, suggesting that French speakers can construe the

 (i) Il a été bête, bête, bête.
 he be.PC.3SG stupid stupid stupid 'He was stupid, stupid, stupid.'

A faithful translation of (i) is difficult to give: it may either have an intensive reading, but potentially also a reading where each stative predicate in the sentence is meant to capture a different state of stupidity manifested through different actions.

¹⁵Verb iteration may be possible with more activity-like states (such as 'sit' or 'stand'), and it may likewise be available for "action-dependent states" like $\hat{e}tre$ $b\hat{e}te$ 'be stupid', e.g. (i) below, but all these predicates pattern with activities for the ongoing-event use (they raise no infelicity in sentences such as (25)).

situation as several drawing events having taken place.¹⁶

(33) Combien de fois Sascha a dessiné? how many of times Sascha draw.PC.3SG 'How many times was Sascha drawing?' (Lit.: 'How many times did Sascha draw?')

Still, the data in (26) remain puzzling, for we do not need to assume that the reported activity took place several times to make these sentences felicitous. Additionally, the difference between French, which accepts ongoing-event interpretations after perfective activities, and English, which does not (see the literal translations of (26)), is left completely unexplained.

4.3 The Notion of Maximality

Atelic predicates combined with the MAX operator as defined in (13) raise another and related issue.¹⁷ To illustrate the problem, suppose that *L* is a stative predicate, for example, $\lambda e.\text{kim-be-sick}(e)$ (see our example (4)). It is generally assumed that stative and activity predicates are cumulative. For *L* to be cumulative means the following: for every event (state) e, e', if L(e) and L(e') each hold, then $L(e \oplus e')$ also holds, where $e \oplus e'$ is the mereological sum of e and e'. Now consider an event (state) e_1 such that $L(e_1)$ holds. Let us ask whether

Unsurprisingly, *une seule fois* 'only once' is clearly an appropriate answer to (33), too. This difference is reminiscent of counting ambiguities in Hungarian, where, depending on the type of the verbal predicate ("ordinary" activity, strictly semelfactive, strictly iterative, ambiguous semelfactive/iterative), counting of minimal events and/or counting of maximal events may be possible or required (see Gyarmathy 2017).

¹⁷We are grateful to Chris Piñón, who drew attention to the problem posed by cumulativity and suggested self-connectedness as a solution (pers. comm.).

¹⁶Note that *être malade* 'be sick' patterns with activities in this respect, differently from the predicates in (25a) and (25b) (cf. *Il a été malade plusieurs fois* 'He was sick several times', vs. *#Il y a eu plusieurs fois un restaurant chinois dans la rue* 'There was several times a Chinese restaurant in the street', acceptable only if there were different times when there was a Chinese restaurant there, which then closed down, then another opened, etc.).

 $MAX(e_1, L)$ also holds according to the definition of MAX (13) and repeated in (34).

- (34) $MAX(e_1, P)$ iff
 - a. e_1 is a part of a possible *P*-event and
 - b. e_1 is not a proper part of any actual event that is part of a possible *P*-event.

Clearly, (34a) is satisfied for P = L because $L(e_1)$ holds: any actual *L*-event is also a possible *L*-event. If e_1 is the biggest actual *L*-event, then (34b) also holds. The problem, however, is that this second condition is unlikely to be satisfied in many realistic scenarios. For example, let us suppose that e_1 is an event (state) in which Kim is sick during time t_1 , that e_2 is an event (state) in which Kim is sick during time t_2 , and that t_1 and t_2 are separated by an interval of time t' during which Kim is not sick. More formally, we basically have the following:

(35) a. $L(e_1) \wedge \tau(e_1) \subseteq t_1$ b. $L(e_2) \wedge \tau(e_2) \subseteq t_2$ c. $\neg \exists e'(L(e') \wedge \tau(e') \subseteq t') \wedge t_1 < t' < t_2$

Since *L* is (by assumption) cumulative, it follows from (35a)–(35b) that $L(e1 \oplus e2)$ also holds. (In other words, the state $e1 \oplus e2$ is also one in which Kim is sick.) Given the scenario described in (35), let us reconsider whether MAX (e_1, L) holds. Again, (34a) is satisfied: e_1 is an actual *L*-event, hence e_1 is also a possible *L*-event. But notice now that (34b) is not satisfied: e_1 is (on the contrary) a proper part of an actual event (namely, $e_1 \oplus e_2$) that is a part of a possible *L*-event ($e_1 \oplus e_2$ is an actual *L*-event, hence also part of an actual *L*-event, hence also part of an actual *L*-event ($e_1 \oplus e_2$ is an actual *L*-event). Consequently, MAX (e_1, L) does not hold in the scenario in (35). (Similarly, MAX (e_2, L) does not hold in this scenario. But if $e_1 \oplus e_2$ is the biggest actual *L*-event, then MAX $(e_1 \oplus e_2, L)$ does hold). Consider now the following French sentence (cf. (4)):

(36) Kim fut/a été malade. Kim be.ps/pc.3sG sick 'Kim was sick.'

Suppose that we represent (36) as we recommended:

(37) $\exists e(L(e) \land \tau(e) \subseteq t_T \land \mathsf{MAX}(e,L))$

Applying the scenario described in (35), imagine that a speaker has e_1 and t_1 in mind when she asserts (36), that is, the event that she asserts to exist is e_1 and the reference time t_T is t_1 . The problem is that, as argued above, the third condition in (37), namely, MAX(e, L), is *false* in the case where $e = e_1$. This is a problem because intuitively, (36) is *true* in the case where the speaker is talking about e_1 and t_1 . (Note that the speaker may not even know about e_2/t_2 .) Analogous scenarios can be applied to other examples of stative or activity predicates (assuming that such predicates are cumulative), so the conclusion is that the definition of MAX adopted in (13) is, in fact, too strong.

We could try to propose another notion of maximality that fares better. One possibility is to use the mereotopological notion of *self-connectedness*. Intuitively, the idea is that an event is self-connected just in case it does not contain any spatiotemporal gaps. (Self-connectedness can be formalized using the notion of connectedness, which in turn is based on the notions of boundary and internal part.) In the scenario described in (35), e_1 and e_2 are (by assumption) each self-connected, but $e_1 \oplus e_2$ is not self-connected, because there is a gap between e_1 and e_2 . The notion *maximal self-connected* (MAX-sc) could be defined as follows (where sc stands for "self-connected" and part(P) for the predicate true of events that are (possibly improper) parts of possible P-events):

- (38) MAX-sc(e, P) iff
 - a. part(P)(e) and
 - b. $sc(e, part(P)) \land$

$$\neg \exists e'(\mathsf{part}(P)(e') \land \mathsf{sc}(e',\mathsf{part}(P)) \land e \sqsubset e')$$

In prose, *e* is maximal self-connected with respect to *P* iff *e* is a part of a possible *P* event (see (34a)), *e* is self-connected, and there is no *e*' such that *e*' is part of a possible *P*-event, *e*' is self-connected, and *e* is a proper part of *e*'. The French sentence in (36) can now be represented as follows:

(39)
$$\exists e(L(e) \land \tau(e) \subseteq t_T \land \mathsf{MAX-sc}(e, L))$$

The formula in (39) is true in the case where the speaker has in mind e_1 for e and t_1 as the value of t_T , which correctly reflects the fact that (36) is intuitively true in this case. In particular, the existence of the later *L*-event e_2 does not render (36)/(39) false, precisely because $e_1 \oplus e_2$ is not self-connected even though it is (as before) an *L*-event.

Note, however, that the definition in (38) helps for states, but it does not draw a distinction between states and activities. So our examples in (26) remain puzzling if MAX-sc is used for activities as well.

One solution is to say what we suggested for these examples in section 4.2: in (present) terms of maximal self-connectedness, the activities described in the first clauses of (26) are (contrary to appearance) really maximal self-connected after all, for they are separated from each other by pauses, albeit very small. When sentences in (26) are taken to describe a single, still ongoing activity, this activity is strictly speaking not self-connected (much like $e_1 \oplus e_2$ with respect to *L* above).

Another solution is to posit a sortal distinction between states and events proper, such that the condition of maximal self-connectedness applies to states but not to events. More precisely, we would extend the typology of perfectives proposed in section 3 as in table 2. Beside strong perfectives (encoding event completion *and* event selfconnected maximality) and weak perfectives (encoding event selfconnected maximality only), we would keep the standard perfective, encoding event completion only. In languages like English, the per-

A Finer-grained Typology of Perfective Operators

| Perfective operator | Requires completion? | Requires maximality? | Semantics |
|---------------------|----------------------|-------------------------|-----------------------------------|
| Weak (Hindi, | No | Yes | [[PFV _M]] |
| Mandarin) | | | |
| Standard (French) | Yes | No | $\llbracket PFV_C \rrbracket$ |
| Strong (French, | Yes | Yes | $\llbracket PFV_{C+M} \rrbracket$ |
| English, Russian) | | | |

 Table 2
 A typology of perfective operators (revised version).

fective is always strong, which accounts for why ongoing-event interpretations after perfectives are infelicitous (recall the literal translations of (26)). But for languages such as French, a strong perfective (encoding completion *and* self-connected maximality) is selected for state predicates, see (40b), and a standard perfective (encoding completion, but not self-connected maximality) is selected for event predicates, see (40a).

(40) a. $\llbracket \mathsf{PFV}_C \rrbracket^{M,g} = \lambda P. \exists e(\tau(e) \subseteq t_T \land P(e))$ b. $\llbracket \mathsf{PFV}_{C+M} \rrbracket^{M,g} = \lambda P. \exists s(\tau(s) \subseteq t_T \land P(s) \land \mathsf{MAX-sc}(s, P))$

The definition given in (40b) is simply the neo-Reichenbachian definition of the perfective (cf. (1)), but now restricted to predicates P of events proper.

It may sound at first sight rather unattractive to posit two different meanings for the very same perfective morphologies in French (the PC and the PS) according to whether they combine with stative or eventive predicates. However, observe that the sortal distinction put aside, (4ob) only adds a condition – the self-connected maximality requirement – to (40a). (40b) is therefore simply a stronger version than (40a). And, in fact, we see an independent reason for positing that the perfective form is semantically stronger when it combines with stative predicates. The relevant observation is that the most obvious alternative to the perfective PS/PC, namely the imperfective form (the *imparfait*) is semantically weaker when combined with stative predicates than when used with eventive predicates. In particu-

lar, when combined with statives, the *imparfait* seems able to express the "event time \subseteq topic time" relation characteristic of the perfective, too. In this perspective, it is less surprising that perfective forms get a specialized, strengthened meaning with stative predicates. We propose that the additional maximality requirement in (40b) with stative predicates reflects exactly this.

To illustrate that the imperfective is semantically weaker when combined with stative predicates than when combined with eventive predicates, let us compare sentences (41a) and (41b).

- (41) a. #La semaine passée, Paul lisait ta lettre. the week last Paul read.IMP.3SG your letter Intended: 'Last week, Paul read your letter.'
 - b. La semaine passée, Paul était malade/triste. the week last Paul be.IMP.3SG sick/sad 'Last week, Paul was sick/sad.'

Out of the blue, (41a) is odd. The reason for this is that with eventive predicates, the *imparfait* has no other choice than expressing the imperfective relation "topic time (*last week*) \subseteq event time (*read your letter*)," which clashes with the assumption that one does not read a letter during a whole week (and note that the problem of (41a) vanishes if we replace the VP by, for instance, *travaillait sur son papier* (work.IMP.3SG on his paper), for it is not unusual to work on a paper during a whole week). This is why such sentences feel incomplete: the reader expects a subsequent clause providing a topic time which could satisfy the aspectual relation "topic time \subseteq event time" without clashing with common assumptions (e.g., … *quand tout à coup*, *son mobile s'est mis à sonner* '… when suddenly, his mobile started ringing').

Let us now look at the stative sentence (41b). Such an imperfective sentence can obviously mean that Paul was sick (or sad) the whole week and perhaps even longer. This corresponds to the imperfective meaning standardly attributed to the *imparfait*, see (42a). Now, imagine that a speaker met Paul for lunch on Monday last week, got
to know that Paul was sick at that time $t_1 \subseteq$ monday, and does not know when Paul recovered exactly, that is, does not know, for any time *t* after t_1 , whether Paul is still sick at *t*. Our observation is that in this scenario, (41b) is *also* true. But note that this second scenario corresponds to the *perfective* meaning (42b), since t_1 is included in Monday, and therefore also included in the topic time provided by the adverbial *last week*.

- (42) a. $\exists e(\operatorname{sick}(e) \land \operatorname{last} \operatorname{week} \subseteq \tau(e) \land \operatorname{theme}(e, \operatorname{paul}))$ (imperfective reading)
 - b. $\exists e(\operatorname{sick}(e) \land \tau(e) \subseteq \operatorname{last week} \land \operatorname{theme}(e, \operatorname{paul}))$ (perfective reading)

Note that (41b) is unacceptable if the speaker knows Paul recovered last week. In other words, with statives, the *imparfait* can express (1) (the standard perfective), but not (40b) (the strong perfective).

That the *imparfait* can convey the aspectual configuration encoded by perfectives when combined with stative predicates is also confirmed by the felicity of the dialogue in (43):

| (43) | a. | La semaine | passée, | tu | étais | n | nalade, | |
|------|----|---|---------|-----|-----------|------|---------|--|
| | | the week | last | you | be.IMP.25 | G si | ick | |
| | | n'est-ce pas? | | | | | | |
| | | QUESTION TAG | | | | | | |
| | | 'Last week you were sick, weren't you?' | | | | | | |
| | b. | Oui. I'ai été | mal | ade | lundi e | et | mardi. | |

yes I be.PC.1SG sick Monday and Tuesday 'Yes, I was sick on Monday and Tuesday.'

The addressee answering (43b) most probably knows when he recovered. If he asserts that he was sick on Monday and Tuesday, he conveys the information he was not sick anymore from Wednesday on. And crucially, this assertion is presented through *oui* 'yes' as a ratification of the truth of (43a). This confirms that the imperfective sentence (43a) can convey the perfective meaning (42b).

In summary, the *imparfait* can only have an imperfective meaning

when combined with eventive predicates. But when combined with stative predicates, the same morphology can additionally convey the same meaning as the standard perfective we have in (1), cf. (42b). We propose that this explains why the perfective morphology, when combined with stative predicates, gets its meaning strengthened and specialized, through the self-connected maximality requirement.

5 Conclusion

At this point, an interesting parallel emerges between the English simple past and the French imparfait: with stative predicates, these two forms can express both imperfectivity and perfectivity. This flexiblity probably reflects the division of labour between aspect markers in these two languages, although in a different way. English has no aspect marker compatible with stative predicates and expressing imperfectivity only (the progressive is typically not acceptable with statives). It is therefore not surprising that the form in charge of expressing perfectivity - the simple past - may also convey imperfectivity with statives. By contrast, French has an imperfective form compatible with statives, but this form has also perfective uses with these predicates, thus competing with the aspectual forms which are 'perfective only' (the PS and the PC) to express perfectivity with statives. It therefore comes as no surprise that the perfective aspect markers – the PS and the PC – get a strengthened perfective meaning when combining with states.

Acknowledgments We are very grateful to Chris Piñón for his careful editing work and his valuable input on section 4.3, as well as Károly Varasdi, Zhang Mengjie and the reviewers of CSSP 2017 and EISS 12 for feedback and discussions. We also thank our informants for their judgments and data: Rajesh Bhatt, Bhamati Dash, Harit Joshi, Ghanshyam Sharma and Yash Sinha (Hindi), Jinhong Liu and Hongyuan Sun (Mandarin Chinese). As far as F. Martin is concerned, this work is supported by the project B5 of the SFB 732 financed by the DFG and hosted by the University of Stuttgart, and by DFG award AL 554/8-1 (Leibniz-Preis 2014) to Artemis Alexiadou.

References

- Altshuler, D. (2014). A typology of partitive aspectual operators. *Natural Language & Linguistic Theory*, 32(3), 735–775.
- Altshuler, D., & Filip, H. (2014). Perfectivity as maximization over event stages. Paper presented at the *Chronos* 11, SNS, Pisa, 16–18 June 2014.
- Bohnemeyer, J. (2014). Aspect vs. relative tense: The case reopened. *Natural Language & Linguistic Theory*, 32(3), 917–954.
- Chief, L. (2008). *Scalarity and incomplete event descriptions in Mandarin Chinese*. Ph.D. thesis, The State University of New York at Buffalo.
- Comrie, B. (1976). Aspect: An Introduction to the Study of Verbal Aspect and Related Problems. Cambridge University Press.
- de Swart, H. (1998). Aspect shift and coercion. *Natural language and Linguistic Theory*, *16*, 347–385.
- Dowty, D. R. (1979). Word Meaning and Montague Grammar. Dordrecht: D. Reidel.
- Filip, H. (1999). *Aspect, Eventuality Types, and Nominal Reference*. New York: Routledge.
- Grønn, A., & von Stechow, A. (2016). Tense. In M. Aloni, & P. Dekker (Eds.) *The Cambridge Handbook of Formal Semantics*, (pp. 313–341). Cambridge University Press.
- Gyarmathy, Zs. (2017). Semelfactives in English and in Hungarian: A countability-based analysis. In N. Brandel (Ed.) *Proceedings of IATL 2016*, vol. 86 of *MIT Working Papers in Linguistics*, (pp. 57–76). Cambridge, MA: MITWPL.
- Higginbotham, J. (2000). On events in linguistic semantics. In J. Higginbotham, F. Pianesi, & A. Varzi (Eds.) Speaking of Events, (pp. 49–79). Oxford: Oxford University Press.
- Klein, W. (1994). Time in Language. London: Routledge.
- Koenig, J.-P., & Muansuwan, N. (2000). How to end without ever finishing: Thai semi-perfectivity. *Journal of Semantics*, 17(2), 147–182.
- Kratzer, A. (1996). Severing the external argument from its verb. In J. Rooryck, & L. Zaring (Eds.) *Phrase Structure and the Lexicon*, (pp. 109–137). Dordrecht: Springer Netherlands.

Landman, F. (1992). The progressive. Natural Language Semantics, 1, 1–32.

Li, C. N., & Thompson, S. A. (1981). *Mandarin Chinese: A Functional Reference Grammar*. Berkeley: University of California Press.

- Lin, J. (2004). Event Structure and the Encoding of Arguments: The Syntax of the Mandarin and English Verb Phrase. Ph.D. thesis, MIT.
- Maisak, T. (1999). Multiple periphrastic perfectives: another case of bounder perfectivization. Ms.

URL http://lingvarium.org/maisak/publ/index-eng.htm

- Martin, F., Demirdache, H., Liu, J., & Sun, H. (2018). Monomorphemic verbs in Mandarin Chinese: Lexical aspect, event structure and non-culminating uses. Ms, Humboldt-Universität zu Berlin and Université de Nantes.
- Partee, B. H. (1973). Some structural analogies beween tenses and pronouns in English. *Journal of Philosophy*, *70*, 601–609.
- Piñón, C. (1997). Achievements in an event semantics. In A. Lawson, & E. Cho (Eds.) *Proceedings of SALT 7*, (pp. 273–296). Ithaca, NY: CLC Publications.
- Piñón, C. (2011). Event structure in event semantics (Aspectual composition with structured denotations). Paper presented at *Chronos 10*, Aston University, Birmingham, 18-20 April 2011.
- Rothstein, S. (2004). *Structuring Events. A Study in the Semantics of Lexical Aspect.* Malden, MA: Blackwell Publishing.
- Schaden, G. (2007). La Sémantique du Parfait. Etude des "Temps Surcomposés" dans un Choix de Langues Germaniques et Romanes. Ph.D. thesis, Paris 8.
- Schaden, G. (2015). An intersective account of localising temporal expressions. In K. Paykin, & C. Majolino (Eds.) *Telling Time*, (pp. 39–71). Cambridge: Cambridge Scholars Publishing.
- Singh, M. (1991). The perfective paradox: Or how to eat your cake and have it too. In L. Sutton, C. Johnson, & R. Shields (Eds.) *Proceedings of the Seventeenth Annual Meeting of the Berkeley Linguistics Society*, (pp. 469–479).
- Singh, M. (1994). *Perfectivity, Definiteness, and Specificity: A Classification of Verbal Predicates in Hindi*. Ph.D. thesis, University of Texas at Austin.
- Smith, C. (1991). The Parameter of Aspect. Dordrecht: Kluwer.
- Smith, C. (1999). Activities: states or events? *Linguistics and Philosophy*, 22, 479–508.
- Taylor, B. (1977). Tense and continuity. Linguistics and Philosophy, 1, 199–220.
- Zucchi, S. (1999). Incomplete events, intensionality and imperfective aspect. *Natural Language Semantics*, 7(2), 179–215.

On (Un)certainty: The Semantic Evolution of Galician *seguramente*

Vítor Míguez

Abstract This paper investigates the semantic development of the Galician adverb *seguramente*. Special attention is paid to the meanings and uses of this item in different periods. The medieval epistemic meaning of certainty contrasts with the present-day meaning of probability. This shift in meaning is explained as a gradual development from certainty to uncertainty with an intermediate stage where the adverb displays features of both certainty and uncertainty markers. It is argued that the rise and fall of manner, strategic, and pragmatic uses correlate with the change in the epistemic use. The ultimate cause for the semantic change of *seguramente* lies in the paradigmatic relations among epistemic markers, and can be better understood if studied in the bigger context of shared discourse traditions.

Keywords adverb \cdot epistemic modality \cdot Galician \cdot semantic change \cdot (un)certainty

V. Míguez, Instituto da Lingua Galega, Universidade de Santiago de Compostela, vitor.miguez@usc.es

In Christopher Pinon (Ed.), Empirical Issues in Syntax and Semantics 12, 217–246. Paris: CSSP. http://www.cssp.cnrs.fr/eiss12/ © 2019 Vitor Miguez

1 Introduction

Seguramente is a prominent epistemic adverb in Galician, a Romance language spoken in Galicia and bordering areas in the northwest of the Iberian Peninsula. This linguistic item is interesting because (i) it underwent a semantic change leading from certainty to uncertainty, and (ii) it provides a benchmark of the reshaping of the linguistic subsystem for epistemic modality during the 20th century. Furthermore, *seguramente* is an instance of *-ment(e)* adverbs, a very productive type of adverbs in Romance languages—analogous to English *-ly* adverbs—of great importance to many linguistic systems, and, in

the particular case of Galician, a key to understanding the present subsystem for epistemic modality.

The reference dictionary for Galician, the *Dicionario da Real Academia Galega* (DRAG) (Real Academia Galega 2012: s.v.), features only one sense for *seguramente*—'with a high degree of probability'. Galician grammars group this adverb together with *quizais* 'maybe' as a device to mitigate an assertion (Freixeiro Mato 2003:160) and with *posiblemente* 'possibly' and *probablemente* 'probably' as part of a system to express different degrees of probability (Álvarez & Xove 2002: 627–628). Some of the examples they offer are the following ones (here and elsewhere, *seguramente* is in italics for perspicuity):

(1) a. Aínda que non o dixese, *seguramente* irá.
'Even though she didn't tell, she will (most) probably go.'

(DRAG: s.v. seguramente)

b. Seguramente o señor concelleiro estará moi ocupado todo o día e non atopará un só intre para vir ata aquí.
'Most probably the city councilor will be too busy the whole day and will not find a moment to come here.'

(Álvarez & Xove 2002:600)

In these examples *seguramente* expresses (high) probability, and, hence, lack of complete certainty. This is especially true for (1a), where *seguramente* could be replaced by *probablemente*, with the former implying a (slightly) higher degree of likelihood. In (1b), uncertainty is exploited to produce a tendentious interpretation which ascribes responsibility to the subject for the negative state of affairs. Such a reading would also arise if *probablemente* were to be used in this context.

The fact that *seguramente* belongs within the paradigm of probability forms can only be surprising if taking into consideration that this adverb derives from *seguro* 'safe, sure', which conveys epistemic certainty—see Vázquez Rozas 2010 for illustration with Spanish data. Nonetheless, it will become obvious from the forthcoming analysis that such a development is not surprising, since there is only a thin line between certainty and uncertainty. Still, the fact that a certainty marker becomes a fully-fledged uncertainty form and loses all certainty semantics deserves further investigation. The aim of this paper is to address this phenomenon through the particular example of Galician *seguramente*.

I will consider epistemic modality a basic semantic category of a scalar nature used to qualify states of affairs (see Nuyts 2001, 2005). Consequently, it will be defined as an estimation of the degree of likelihood that an event takes/has taken/will take place, as made by the speaker. Several positions may be distinguished on the epistemic scale, ranging from positive to negative certainty, with intermediary uncertainty values in between. The case that will be addressed here involves a demotion on the scale, leading from certainty to the uncertainty value of probability.

Corpus data will be used as the main source of evidence. I will resort primarily to the *Tesouro informatizado da lingua galega* (TILG) (Santamarina 2014), which contains over 26 million words written between 1612 and 2013. It currently includes 1070 distinct occurrences of *seguramente*. For medieval data I will use the *Tesouro medieval informatizado da lingua galega* (TMILG) (Varela Barreiro 2007), consisting of 16000 textual units corresponding to the period c. 1200–1600. 23 different instances of the adverb are found in this corpus. Both TILG and TMILG were developed at the Instituto da Lingua Galega. I will consider three chronological periods: medieval (1200–1500), early contemporary (1880–1930), and present day (1975–2013).

This paper is organized as follows. Section 2 reviews previous work on (un)certainty adverbials. Section 3 presents the data, showing the different ways in which the adverb under consideration is and was used. Section 4 provides an interpretation of the diachronic path of *seguramente*. Section 5 adds some discussion and concludes.

2 Markers of (Un)certainty

The default expression for certainty in language is a bare assertion. Therefore, the use of marked expressions of certainty must be for a compelling reason. In fact, it is commonly accepted that this part of the epistemic modal system lies in a paradox: "the fact that we only say we are certain when we are not" (Halliday 2004:625). Sometimes speakers/writers are fully aware of this:

(2) E, obviamente, nos territorios da incerteza antes referidos, en que, máis unha vez, nos moveremos, non poderei evitar os «tal vez», os «quizais», os «segura-», os «probábel-» ou os «posibelmente», e mais os «sen dúbida», os «con certeza» ou os «evidentemente», que, de maneira directa ou paradoxal, sosteñen as precarias reconstrucións e as salutares dúbidas (...) (1999, MPRIND999, 78, TILG)

'And, obviously, in the aforementioned territories of uncertainty, in which more than once we will be moving, I will not be able to avoid «tal vez» ['perhaps'], «quizais» ['maybe'], «seguramente» ['surely'], «probabelmente» ['probably'] or «posibelmente» ['possibly'], nor «sen dúbida» ['no doubt'], «con certeza» ['for sure'] or «evidentemente» ['evidently'], which directly or paradoxically support the precarious reconstructions and the salutary doubts.'

Crucially, the author of this excerpt—from a study on Galician medieval literature—divides the inhabitants of the land of uncertainty into two groups: those directly pointing to doubt (possibility and probability adverbials), and those doing so paradoxically (certainty adverbials). Interestingly enough, *seguramente* is included in the first group.

Expressions of certainty signal that certainty cannot be taken for granted. They are exploited rhetorically to confront any shadow of doubt that may concern the state of affairs they introduce, and, thus, they lead to inferences of uncertainty (cf. Aijmer 2002, Simon-Vandenbergen 2007, Simon-Vandenbergen & Aijmer 2007). This is, in essence, the paradox of certainty expressions.

A good illustration of how certainty expressions work is found in *certainly*. According to Aijmer (2002), *certainly* is used as an emphasizing device, either in negative contexts or to reinforce agreement, and can also serve a contrastive function. Byloo et al. (2007) found

that the emphasizing, or strengthening (the term used here, see section 3.3), function is absolutely dominant for *certainly*, as compared to the epistemic function: according to their minimal count, the epistemic use of *certainly* accounts for around 8% of cases, whereas the strengthening use represents more than 40%.

In contrast with certainty markers, epistemic expressions of uncertainty, like *probably* and *maybe*, are not paradoxical. They qualify a state of affairs as (im)probable or possible, and, when used strategically,¹ they mitigate illocutionary force. In this connection, uncertainty markers are the mirror image of expressions of certainty.

There is, however, a third type of markers, considering both their epistemic and non-epistemic functions. Byloo et al. (2007) compared English *certainly* with Dutch *zeker*, and found that the epistemic meaning was infrequent and differed in both expressions. Both adverbs express certainty, but only *zeker* can also convey a weaker value of (high) probability.

A good deal of attention has been devoted to English adverbial forms *no doubt* and *surely* (Aijmer 2002, Carretero 2012, Downing 2001, 2008, Simon-Vandenbergen 2007, Simon-Vandenbergen & Aijmer 2007, Traugott 2014). Like Dutch *zeker*, these epistemic markers are characterized by semantic variability: they convey either certainty or uncertainty depending on the context. When *no doubt* cooccurs with the auxiliaries *will* and *would*, it "clearly expresses the speakers' conviction that the state of affairs took or will take place but that they are not in a position to have absolute certainty" (Simon-Vandenbergen 2007:15). This use is in line with the paradoxical nature of certainty markers. *No doubt* also functions as an expression of certainty when preceding a contrastive clause with *but*, in which it has a concessive meaning, just like *certainly* (Simon-Vandenbergen & Aijmer 2007:95, 131). However, *no doubt* can also express probability, as evidenced by translation equivalents: it may correspond to am-

¹Following Nuyts 2001, I refer to strengthening and mitigation as *strategic uses*, since their occurrence depends on factors aimed at achieving special effects in discourse, that is, they are the result of a discourse strategy.

bivalent expressions like Dutch *zeker* or Swedish *säkert*, but also to forms coding probability, like Dutch *waarschijnlijk*, Swedish *nog*, and French *il est probable* (Simon-Vandenbergen & Aijmer 2007:129–132).

Surely, like *no doubt*, developed a contrastive function and usually collocates with *but*. Like *certainly*, but unlike *no doubt*, *surely* features strengthening uses in which, rather than expressing epistemic certainty, it emphasizes the truth of the utterance (Downing 2001, Simon-Vandenbergen & Aijmer 2007:137–139, Traugott 2014). Nevertheless, there is agreement that *surely* has become dissociated, in several contexts, from the meaning of assurance it overtly points to. In the "confirmation-seeking" function, for instance, *surely* expresses probability and invites the hearer to agree (Aijmer 2002, Simon-Vandenbergen & Aijmer 2007, Traugott 2014). Furthermore, this adverb features a variety of other functions, ranging from inferential evidentiality and mirativity to challenge and persuasion (Downing 2001, Carretero 2012), which make it an interesting object of inquiry.

Given the previous considerations, we can conclude that there are limits to the ways epistemic and strategic uses combine within a linguistic unit. Those units conveying an epistemic value of certainty may be used to reinforce an assertion (strengthening), whereas those coding values of probability or possibility may be used to lower the strength of the utterance (mitigation). There are also linguistic devices whose epistemic value varies depending on the context, and that exhibit a wide range of non-epistemic functions, one of them usually being strengthening, but not mitigation.

3 Data and Uses of seguramente

The history of the Galician language is characterized by different periods of splendor and darkness, which directly determine to what extent the language was used for writing and, consequently, how much linguistic data from past eras have reached us. Galician was the main written and spoken language in Galicia during the late Middle Ages (c. 1200–1500), but after 1500 three centuries of darkness follow in which the written form of the language is almost completely aban-

doned. As a consequence, little is known of the modern language in its early stages—see Mariño Paz 2008 for details on the history of Galician.

Adverbs in -mente are scarce in medieval documents, at least as compared to present-day language. Nevertheless, they feature interesting characteristics which signal future paths of development. According to TILG, (epistemic) -mente adverbs are rare in the modern language until the 1880s. Since the corpus at this point was of considerable size, their (re)emergence can be linked to the functional expansion of the language: prose, the new text type, requires "new" epistemic and discourse markers. Thus, the emergence of -mente adverbs in modern Galician can be taken to signal the transition from a "rural" written language, used primarily for folk-like poetry, to a "learned" written language, used for different types of prose (fictional, journalistic, essayistic, technical). It is significant that the main certainty marker in TILG from the 18th century to the late 19th century is abofé (from Latin AD BONA FIDE 'in good faith'), today regarded as archaic, and that this item decreases in frequency from 1900 onward. It is at this point that other, more nuanced expressions of certainty arise, namely -mente adverbs like certamente 'certainly', realmente 'really, actually', seguramente 'surely', and verdadeiramente 'truly'. In present-day Galician, these items play a central role in the expression of certainty and strengthening.

In the following, I will illustrate the different uses *seguramente* could and can be put to. This will set the basis for the interpretation of its evolution. Ideally, this will show that, despite the empirical limitations, an interpretation of the diachronic path of this adverb is of interest to current research on language change.

3.1 Manner Uses

In Latin, an adjective modifying the feminine noun MENS, MENTIS 'mind, mood' in the ablative case was used to convey a manner of action. In Romance languages, the ablative form MENTE led to a fully-fledged derivative suffix which creates adverbs out of adjectives in

their feminine form. This mechanism was already operative in the medieval period of the Galician language (see Ferreiro 2001:206–207). Thus, *seguramente* derives from the feminine form of the adjective *seguro* 'safe, secure'.

Manner uses of *seguramente* account for half of the data in the medieval period. In these cases, the adverb refers to how the action is carried out. This may pertain to different meanings, such as 'safely' (3a), 'confidently' (3b), or 'steadfastly' (3c), which are reminiscent of manner uses of English *surely* (see Downing 2008, Traugott 2014).

- (3) a. Et sabede que en aquel tépo era costume que todo messageyro andasse en saluo per huquer, et que segurament cõtasse seu messagê, et nũca por ende mal rreçebessen. (1370–1373, CT 20/236, TMILG)
 'Be it known unto you that at that time it was customary for a messenger to safely walk everywhere, and that he safely told his message, and that no harm was done to him because of that.'
 - b. Sal ja da arca *seguramẽte* tu et tua moller et teus fillos et suas molleres contigo, et todaslas anymalias que convosco forõ ena arca; (1300–1330, XH II/8, TMILG)

'Now get out of the ark confidently you and your wife and your sons and their wives with you, and every animal that was with you in the ark.'

c. Et mandoos que laurassem et criassem *seguramente* et que llj dessem seu peyto, segũdo que o dauã a seu rrey. (1295–1312, TC 346/512, TMILG)

'And he ordered them to work and breed steadfastly and to pay him a tax, as they did with their king.'

In present-day language, manner uses like the ones in (4) constitute less than 1% of all occurrences. They are very rare, but prove that the manner use is still accessible for contemporary speakers/writers. Since the suffix *-mente* is a fully productive mechanism to create adverbs in present-day Galician, *seguramente* can be used as the manner adverb for *seguro*.

- (4) a. ... mesmo se non pasa de especie fraudulenta a afirmación idealista de que a URSS desenvolveríase tan rápida e *seguramente* que de contado había ultrapasar en tódo-los terreos ós Estados Unidos ... (1980, MEFCRO980 26, TILG)
 '... even if merely deceitful, the idealist affirmation that the USSR would develop as fast and securely/steadfastly that it would immediately surpass the United States at all levels ...'
 - b. Mais, polo que toca á súa preocupación polo modo no que poden estar seguros do seu estado de salvación e de graza, será máis *seguramente* obtido dos nosos ... libros ca dos escribáns ingleses. (2006, WBRETN006 166, TILG)

'But, regarding your concern about the way you can be sure of your state of salvation and grace, it will be more securely obtained from our books than from the English scribes.'

3.2 Epistemic Uses

Epistemic uses are prone to ambiguity, at least in two ways: they may be ambiguous (i) with other uses, namely, strategic ones (strengthening and mitigation), and (ii) between several epistemic values, namely, certainty and (high) probability (Byloo et al. 2007, Simon-Vandenbergen 2007). In this section, I intend to show that *seguramente* conveys different epistemic values in different periods.

In the Middle Ages, *seguramente* was used to express epistemic modality from the earliest attestations. This is in contrast to *certa-mente* and *realmente*, which were used as manner adverbs at that time.

(5) E por que esto que dizian non era mui sen razon, ca d'aver ela seu fillo estava ena sazon; e avia tan gran fever, que quena viya enton dizia: "Seguramente, desta non escapará". (1264–1284, CSM 256/26, TMILG)

'And they said so not without reason, because she was about to have her child; and she had such a big fever, that those who saw her at that moment said: "Surely, she will not get out of this".' (5) is a stanza from the Cantigas de Santa Maria, a literary work consisting of poems focusing on the Virgin Mary. Many of them, like the present one, tell a story in which the Virgin performs a miracle. Although for the modern reader the probability meaning of seguramente is available in (5), it is very unlikely that this was the meaning coded by ancient speakers/writers. First, the adverb expresses the certainty of an imminent death, which is the prelude for the miracle to take place. Of course, the writer knew this, but "those who saw her" did not know about the coming miracle. In any case, the writer made sure that the sick woman looked really bad: in (5) we learn that she was about to give birth and that she had a terrible fever, and in the preceding context some physicians say that she will not live long. In short, the context does not allow for a shadow of doubt, so the use of a probability expression would be odd. Second, the most suitable diachronic path for seguramente is one where the manner function leads to strengthening and epistemic certainty uses, and the Middle Ages represent the early stages of this process (see section 4.1).

In present-day Galician, examples like (2) clearly show that *seguramente* is seen as an expression of uncertainty. Indeed, it is used to convey (high) probability. This is especially frequent in essayistic prose, where writers tend to assess the likelihood of a proposed explanation, as in (6).

(6) a. Costume aínda practicado de deixar pan e o lume aceso na noite de Nadal, para as visitas que nos fan as ánimas dos nosos parentes mortos. *Seguramente* é unha pervivencia dos rituais funerarios prehistóricos e romanos. (1999, CUADIC999 86, TILG)

> 'Tradition still in force of leaving bread and the fire burning on Christmas Eve, for the visits that will be paid to us by the souls of our dead relatives. This is (very) probably a preservation of prehistorical and Roman funerary rituals.'

b. Carlos Rico, un antigo emigrante en Bos Aires[,] aprende alá o oficio, *seguramente* por influencia italiana. (2008, LLRFER008 68, TILG)
 (Carlos Rico, a former expatriate in Buones Aires, learns the trade

'Carlos Rico, a former expatriate in Buenos Aires, learns the trade there, (very) probably by Italian influence.'

Early contemporary examples of epistemic *seguramente*, like those in (7), represent a middle ground between the two other stages.

- (7) a. Dios lle conserve a vida ó siñor Pondal e lle dea saúde pra que, canto antes, poidamos saborear súas novas produciós, que *seguramente* serán unha notabre adquisición prá literatura rexional. (1886, RODFIL886 119, TILG)
 'May God preserve the life of mister Pondal and give him health so that, as soon as possible, we can relish his new works, which will surely/no doubt/?(very) probably be a remarkable acquisition for the regional literature.'
 - b. Dáme o corpo que quen fixo tal escamoteo é, seguramente, partidario de qu'as imaxes se arromben[,] nin máis nin menos, que asegún foron. (1886, GAL076886 4, TILG)
 'I have a feeling that the person who did this legerdemain is surely/no doubt/(very) probably in favor of arranging the statues nothing less than as they were.'

(7a), with the modalized verb in the future tense, strongly reminds one of the collocations of *no doubt* with *will* and *would* (see section 2). In such contexts, the epistemic item conveys the speaker's conviction that the state of affairs will or would apply (i.e., certainty), even though this certainty is weakened by the association of the situation with futurity—see Dahl 2000 on the different grounds for talking about the future. Nevertheless, a meaning of uncertainty does not seem to fit in with the context, because it would conflict with the (flatterer) first part of the utterance.

The meaning of *seguramente* in (7b) seems closer to probability than to certainty, especially because of the influence of the opening expression pointing to the realm of intuition. Moreover, the broader context of this example, which cannot be reproduced here for the sake of space, is clearly ironic and reminiscent of the tendentious interpretation achieved in (1b) by means of expressions of probability.

3.3 Strategic Uses

Strategic uses refer to a modification of illocutionary force rather than a qualification of a state of affairs. The use of linguistic expressions in this fashion is due to pragmatic factors: the need to reinforce an assertion has to do with the speaker's rhetorical purposes (show agreement with the interlocutor or push one's ideas), whereas the motivation to mitigate an utterance is usually related to a facesaving strategy (regarding either the speaker's or the hearer's public image)—see, for example, Brown & Levinson 1987, Caffi 1999.

In accordance with its epistemic meaning of certainty, *seguramente* was used as a strengthening device during the Middle Ages.

(8) Et cõ todo esto era tã ben feyto eno corpo et eno rrostro que nõ achariades nehũ tãto, et *segurament*, como diz Dayres, ben pareçía caualleyro estando en praça. (1372–1373, CT 75/274, TMILG) 'And because all of this he had such a well-made body and face that you would not find another like him, and, indeed, as Dares says, he looked like a knight standing in the battlefield.'

(8) is part of a description of Hector, the Trojan prince, where the writer glorifies his physical and moral qualities. Here, *segurament(e)* reinforces the writer's aesthetic judgment, which, in turn, is based on an external source. Thus, *segurament(e)* is used by the writer to underscore his assessment, but also to show agreement with a third party.

Clear cases of strengthening uses like (8) are not the norm. In fact, strategic uses in general are very difficult to tell apart from epistemic uses:

On (Un)certainty: The Semantic Evolution of Galician seguramente

(9) Vedes, fremosa mia senhor, *segurament(e)* o que farei: Entanto com'eu vivo for, nunca vus mia coita direi; ca non m'avedes a creer, macar me vejades morrer. (1220–1240, LP 031/327, TMILG) 'See, my beautiful lady, what I will surely do: as long as I am alive, I will never tell you my grief; since you would not believe me, even if you saw me die.'

The male voice in (9), a fragment of a love poem, is devastated by the indifference of the lady he loves. *Seguramente* in this case can be read as a strengthening marker: in a context of bold determination such as this one, the writer uses the adverb to express that he means what he says, that he will of course honor his word and refrain from conveying his passion and grief. However, it also makes sense for the writer to qualify his own intentions epistemically: in this reading, he estimates as 100% likely that he will refrain from sharing his feelings, since the lady will not believe him, no matter what happens. This kind of ambiguity is quite frequent, and it represents a cornerstone for understanding the diachronic relations between meanings (see section 4).

The examples in (6) are clearly purely epistemic, and so is the vast majority of examples in TILG for present-day Galician. The idea, present in Caffi 1999, that any instance of an expression of uncertainty constitutes a case of mitigation is not adequate, "since that would leave the notion of strategy void of meaning" (Nuyts 2001:101). An example in my data that can tentatively be interpreted as mitigating is the following one:

(10) Cando nota a cara seria e incrédula dela descúlpase. Vós na cidade seguramente tendes outras ideas, non vos podedes preocupar dos alumnos fóra das horas de clase. (1989, HEIANA989 66, TILG)

'When he notices her serious and skeptical face, he apologizes: "In the city you probably have other ideas, you cannot worry about students outside of school hours".'

(10) features seguramente in a context where the speaker directly ad-

dresses his interlocutor. The use of *seguramente* here can be seen as an attempt by the speaker to mitigate a statement that assigns a particular view to the hearer, which in itself is problematic. Still, it is difficult to tell whether this is a real strategic use or just a byproduct of the epistemic meaning.

Early contemporary examples of strategic uses are scarce but show clearly that the strengthening use was still in force at the time.

(11) Seguramente que é un absurdo que un rapaz galego non poda chegar a ser arquitecto na súa terra quedando asín a carreira d'arquitectura reservada non aos mellores senón aos que poidan té-los medios económicos. (1922, REX004922 5, TILG)
'It is surely nonsense that a Galician boy cannot become an architect in his homeland, thus resulting in the situation that the architecture career is reserved not for the best but for those who can afford it.'

(11) is part of an opinion article where the writer expresses some critical thoughts on the university situation of Galicia at the time. In that context, an epistemic qualification of the nonsense of a state of affairs is odd, as compared to a strengthening use through which the writer underscores her or his position.

3.4 Pragmatic Uses

Pragmatic uses are identified for Dutch *zeker* by Byloo et al. (2007) as cases where the adverb modifies the nature of the speech act, particularly, turning a declarative into a special kind of interrogative what they call a "declarogative." The speaker uses this mechanism to ask for confirmation of her claims. This matches perfectly what Simon-Vandenbergen & Aijmer (2007:136), discussing *surely*, term "the opening-up function, reaching out to the addressee for confirmation." Some of the examples they offer are these:

- (12) a. Hij zal wel weten wat ie kan *zeker*?'He'll know what he can do, won't he?' (Byloo et al. 2007:52)
 - b. Of course the chaos when the Supreme Being was discovered tied up and concussed on the floor would be indescribable, but

surely they would need to be more than just lucky to win much more time out of mere chaos?

(Simon-Vandenbergen & Aijmer 2007:136)

A good indicator of the unique nature of these interrogatives is the fact, as pointed out by Byloo et al. (2007), that they retain declarative syntax. They are not real questions and the adverb does not function as an epistemic or strengthening device, but as a signal of this particular kind of speech act. In my data, there is a single instance of *seguramente* used in this way, corresponding to the early contemporary period.

Mar.:-¿Xa pensas botarlle o lazo? Mer.:-Como poida (13) atrapá-lo... Non é fácil atopar un bon partido non saíndo d'aquí, de sorte que se algún aparece por casualidá, hai que tratar de cazá-lo... Inda que teña que correr atrás del, como cando andamos perseguindo as bolboretas. Mar.:--¿E seguramente irá hoxe ao baile do Casino? Mer .:-- Por eso vou eu. Mar .:- Daquela é mellor para o éisito dos teus proieitos que m'eu quede na casa (rindo). (1920, CREPEC920 10, TILG) 'Mar .:- Are you already thinking of getting him? Mer .:- If I could catch him... It isn't easy to find a good catch without getting out of here, so if someone appears by chance, one must try to catch him... Even if you have to run after him, like when chasing butterflies. Mar .:-- And surely he will go today to the ball at the Casino? Mer .:- That's why I'm going. Mar .:- Then it will be best for the success of your plan that I stay at home [laughing].'

The use of *seguramente* in (13) is a clear case of a declarogative: Mar. seeks confirmation of a fact that she assumes to be the case. It is significant that Mer.'s reply does not answer the question but elaborates on the topic, thus treating Mar.'s question as a declarative clause.

4 Development and Change of the Epistemic Function

(Inter)subjectification is a highly influential notion in diachronic language change (Traugott & Dasher 2002, Traugott 2010; see also LópezCouso 2010 and Nuyts 2014). According to it, linguistic forms strongly tend to evolve from objective to subjective and intersubjective. In other words, meanings develop from the description of the world (objective) to the speaker's attitudes and beliefs towards the world (subjective) and the speaker's stance towards the hearer (intersubjective). In the context of the systems of qualifications of states of affairs, subjectification can be seen as a widening of the perspective on the event (e.g., the development of deontic meanings out of dynamic modality) and intersubjectification as the recruitment of a linguistic item into the field of interaction management and out of the system of qualifications (e.g., the shift from epistemic towards discourse marker)—see Nuyts 2014 for details.

Over the last decades, many advances have been made in understanding semantic change, and adverbials have been at the heart of the discussion. A well-known pattern is that meanings tend to become more (inter)subjective while structurally increasing the scope of the expression (see Lenker 2010, Traugott & Dasher 2002). This is represented in the clines in (14).

(14) a. non-subjective > subjective > intersubjective

```
(Traugott & Dasher 2002:281)
```

 b. clause-internal adverbial > sentence adverbial > discourse particle (Lenker 2010:117)

According to Downing (2008:679), English *surely* developed from a manner adverb, through an epistemic marker, to "an opaque, speakeroriented indexical," in accordance with these ideas. One could suppose that the evolution of *seguramente* follows this same diachronic trajectory. However, this does not seem to be the case: the "weakening" of the epistemic meaning is a shift from one type of subjective meaning to another type of subjective meaning; and it is not clear that the epistemic use is older than the strengthening use, or that there is a more plausible path from manner to epistemic modality than from manner to strengthening. Thus, if we conclude that certainty led to probability, and that strengthening led to epistemic modality, the evolution of *seguramente* would be partially alien to (inter)subjectification, and partially contrary to it. In the next sections, I will try to show that this is, indeed, the case. After some considerations on the development of the "post-manner" functions in section 4.1, I address the semantic change of this adverb in section 4.2, and add some additional evidence in section 4.3.

4.1 From Manner Adverb to Discourse and Modal Marker

The available medieval data offer no clear proof for the claim that the manner use preceded the epistemic or the strengthening use: the three meanings are present from the earliest moment. Nevertheless, some factors strongly suggest that the manner use is older, and that it led to non-manner meanings. First, as stated in section 3.1, there was a construction with MENTE in Latin that already conveyed manner of action. Second, we observe an overall tendency for the manner use to disappear, with a radical contraction from more than 50% in the Middle Ages to less than 1% in the whole contemporary period. Other present-day epistemic and discourse markers, like *certamente* and *realmente*, also show this tendency. Finally, there exists a suitable semantic path from the manner meanings of the adverb to the strengthening use, but it seems difficult to reach manner meanings from either strengthening or epistemic modality.

Traugott (2014:79) argues that the meanings 'carefully' and 'steadfastly' led to implicatures of truthfulness in the case of *surely*, and that these implicatures became conventionalized, in particular with verbs of locution and hearing. In the case of *seguramente*, the meanings 'safely' and 'confidently' may have led to the same implicatures in similar contexts:

(15) E por ende disso hũ sabio que ouo nome Tulio, que en nehũa cousa nõ pode homẽ auer amĩgo a quẽ podesse dizer seguramente sua uoentade, assi co mo a ssi meesmo; (1300–1350, CPa XXVII, 4/126, TMILG)

'Thereby a wise man named Tullius said that no one in any respect can have a friend to whom he can confidently/truthfully tell his will, as he does to himself.'

According to Traugott (2014:88), the manner adverb with the meaning 'in a truthful manner' led to an epistemic marker. However, she does not make a clear difference between strengthening and epistemic uses, so it is difficult to know which one was first. Interestingly, she explains that in the older period *surely* "has mainly an emphasizing function, foregrounding the speaker/author's stance toward the content of the clause" (Traugott 2014:82). Therefore, it is not unlikely that *surely* was first a strengthening device, and only later an epistemic marker.

Let's pay attention to ambiguities, that is, the presence of several meanings in a particular use of a linguistic item. Ambiguities are important, since they may reveal the presence of invited inferences, which may lead to the conventionalization of a new meaning (Traugott & Dasher 2002). If we take a look at older occurrences of *seguramente*, we find ambiguities between manner and strengthening (16), and also between strengthening and epistemic modality (recall (9)), but not between manner and epistemic modality.

(16) Et ela quando descobrio os nébros do moço, forõ eles tã espãtados, que nõ poderõ falar. Et ela lles diso: —Meus fillos, comede *seguramête* que meu fillo he, et o trouxe eno ventre et ja del comĩ ... (1390–1420, MS III, 2/43, TMILG)
'And when she uncovered the limbs of the kid, they were so scared that they could not speak. And she told them: —My children, eat with confidence since he is my child/eat since he is indeed my child, and I carried him in my womb and I already ate him.'

Manner uses related to safety, confidence, steadfastness, or truth eas-

ily lead to inferences of reinforcement of the whole state of affairs: if someone walks safely, works steadfastly, or speaks truthfully the adverb contributes to strengthening the fact that the event takes place. Moreover, epistemic modality does not seem compatible with such manner uses, since the former implies the questioning of the factuality of the state of affairs, which the latter entails. As the implicated meaning of reinforcement becomes more frequent and spreads across contexts, the adverb gets positional freedom, thus developing sentential syntax.

The transition from strengthening to epistemic modality is not a radical one. Once the adverb scopes over the clause reinforcing its content, it may be employed in contexts where (it is known that) the speaker does not (or cannot) have absolute certainty about the factuality of the state of affairs, like (9) and (17), which are ambiguous between a strengthening and an epistemic reading.

(17) Et mays te digo; nõ era Troylos vilão que fose asi rroubado nẽ perdido nẽ prendido de nẽgũ ca ben sey que ẽno mũdo nõ ha mellor caualeyro que el et *seguramẽte* que el querra seer ben entregado d'esta prenda. (1350–1399, HT 199/192, TMILG)
'And I will tell you more; Troilus was not a villain that could be robbed, nor lost, nor captured like that by anyone, since I well know that there is no better knight than him in the world, and he will surely want to be satisfied like this.'

4.2 From Certainty to Uncertainty

From a diachronic point of view, it is fair to assume that certainty preceded, and, thus, led to, uncertainty in cases like those of *zeker*, *no doubt*, and *surely*, presented in section 2. The aforementioned paradox of certainty expressions provides the enabling context for the change to take place: since qualifying a state of affairs as certain is semantically redundant, doing so gives rise to inferences that the statement is disputed, and that there may be some reason to doubt it. When such an inference conventionalizes, the epistemic meaning of certainty becomes inaccessible, and a new marker of uncertainty

emerges. Given the continuous nature of a large part of semantic change (Traugott & Dasher 2002), the logical next step after certainty would be a high value of probability: in contrast to possibility, probability is a gradual category (witness the ability of probability expressions to be modified by degree adverbs), and it is closer to certainty in that it implies a stronger commitment to the state of affairs, as evidenced in (18).

| (18) | a. | Maybe she is at home, but I don't think so. | [possibility] |
|------|----|--|---------------|
| | b. | She is probably at home, #but I don't think so. | [probability] |
| | c. | She is certainly at home, #but I don't think so. | [certainty] |

I will contend that the semantic evolution of *seguramente* followed a pathway from certainty to (high) probability, and that the nonepistemic uses present at the different stages of the process further support this claim.

Two main types of strategic uses have been distinguished: strengthening and mitigating. The latter is linked with uncertainty expressions, whereas the former is associated with expressions of certainty and also with ambivalent expressions like *zeker*, *no doubt*, and *surely* (see section 2). *Seguramente* features the strengthening use during the Middle Ages and the contemporary period until at least the 1920s see (11). This means that during that time the adverb functioned as an expression of either certainty or both certainty and uncertainty. As pointed out in section 3.2, during the medieval era it coded certainty, while in the early contemporary language it expressed either certainty or uncertainty, depending on the context.

Byloo et al. (2007) relate the pragmatic use of *zeker* to its weaker epistemic meaning. They also point out that *zeker* does but *certainly* does not feature this use. Of course, epistemically ambivalent expressions are a very tight fit to the hybrid clause type declarogatives represent: the speaker seeks confirmation from the hearer but at the same time wants to assert her near certainty. Interestingly enough, the pragmatic use is not available for certainty and uncertainty items. It seems, in fact, to be exclusive of expressions like *zeker* and *surely*. Given the considerations above, it seems safe to claim that epistemic *seguramente* arose as a marker of certainty during the Middle Ages, and became weaker over time. Considering the available data for Galician, it is impossible to know precisely when *seguramente* started to behave as an epistemically ambivalent expression —but see section 4.3. What we do know is that as late as in the 1920s *seguramente* had some properties that made it different from its current status as a pure marker of probability. This is what the latest strengthening and pragmatic uses are telling us.

The question remains as to why seguramente changed in the first place. The answer may have to do with the principle of 'no synonymy' (e.g., Croft 2000:176-178), which predicts that situations of synonymy tend to be avoided by speakers. This is probably what happened with the paradigm of strengthening and epistemic -mente adverbs in Galician, and presumably in other Romance languages as well. Many of these items evolved from manner adverbs to discourse and modal markers, for instance certamente, efectivamente, realmente, seguramente (see Rivas & Sánchez-Ayala 2012 and Villar Díaz 2013 for data on Spanish). Some of them specialized in certainty and strengthening uses (certamente), others developed different discourse functions (efectivamente, realmente), whereas others had to go down the epistemic scale to find their place in the paradigm of epistemic expressions (seguramente). One must still account for why seguramente developed further in the direction of weak epistemic meaning, instead of stabilizing as a weakened certainty item, as surely has presumably done. The emergence of the adverbial uses of the adjective seguro is likely to be connected with this, since the latter displays the typical characteristics of epistemically ambivalent forms and, furthermore, is lexically related to seguramente (see Company Company 2017, Vázquez Rozas 2010). The final question would be, then, how seguramente deviates from probablemente as to be kept as a distinct probability marker. I can only suggest here that seguramente must code a (slightly) higher probability value, and be a subjective (in the sense of Nuyts 2001) epistemic adverb.

4.3 Semantic Change, Language Contact, and Discourse Traditions

Thus far, this paper has discussed data from Galician, an understudied language which presents drawbacks for historical investigation. Both Galician and Spanish are official in Galicia, and Galician speakers usually have a good command of Spanish and use it to differing degrees. For the case at issue, it would be difficult to maintain that Galician-Spanish bilingual speakers have separate semantic representations for *seguramente* in each language: the functions of the adverb in the two languages are too close. Conversely, it would be too simple to argue that a change generated in one of the languages, namely Spanish, the hegemonic language, and was adopted by the other. The situation seems more complex.

Hummel (2013) studies the diachronic expansion of -ment(e) adverbs in Romance languages and warns about the prevalent habit of operating under the logic of national languages with loanwords. This usually leads to ignoring the continuity between Latin and Romance (learned) spoken and written traditions, and to overlook the influence of the shared practice of writing Latin on the development of written Romance and English-no doubt one should also include other European languages on the list. Hummel (2013) highlights the fact that the English equivalents of the 10 most common -mente adverbs in European Spanish are also widely used. This is assumed to prove the cross-cultural nature of -ment(e) (and -ly) adverbs and their development within a long shared culture. Of special interest for the present study is the idea that some particular historical developments can also be common to a wide number of languages: according to Hummel, the discourse functions of *-ment(e)* adverbs in the main European languages arose in the 19th century and became general in the 20th century.

When looking at (the change in) the use of particular units, it may be sometimes necessary to pinpoint a more specific context than a general European linguistic tradition. The development of *segurament(e)* as a pure marker of probability appears to be a "Spanish phenomenon," that is, a feature telling apart languages spoken in Spain from neighboring languages.

- (19) Epistemic meanings of *segurament(e)* and cognates in the main languages of the Iberian Peninsula²
 - a. Asturian: *de manera cuasi segura* 'almost certainly'. (DALLA: s.v. *seguramente*)
 - b. Basque: *segur aski*, used to express what you think will be the case. (EH: s.v. *segur*)
 - c. Catalan: probablement 'probably'. (DIEC: s.v. segurament)
 - d. Galician: with a high degree of probability. (DRAG: s.v. *segura-mente*)
 - e. Spanish: *probablemente, acaso 'probably, perhaps'*. (DRAE: s.v. *se-guramente*)
 - f. Portuguese: with great certainty; *certamente, decerto* 'certainly, by all means'. (*Dicionário Priberam da Língua Portuguesa*: s.v. *seguramente*)

Lexical entries³ (19a) to (19e), corresponding to the languages of Spain, link *segurament(e)* to different degrees of uncertainty, ranging from high probability to possibility. Even Basque, a genetically unrelated language, has a form with the same Latin lexical stem and similar meaning. The entry for Portuguese (19f), by contrast, clearly links *seguramente* to certainty.

Nevertheless, dictionary entries should not be taken as hard proof for this claim, since they usually disregard the polyfunctionality of epistemic and discourse markers. What is really interesting is the fact that none of the lexical items (19a) to (19e) allows strengthening uses, whereas they are common in Portuguese, for example, (20).

²DALLA = Academia de la Llingua Asturiana (2015), EH = Euskaltzaindia (2016), DIEC = Institut d'Estudis Catalans (2007), DRAE = Real Academia Española (2014). All online dictionary entries were retrieved on 03-28-2018.

³The resort to dictionaries responds to practical criteria: there are no better sources of semantic evidence for most of the items considered, and they offer comparable information.

(20) Há de haver um código de indumentária e acho isso aceitável, mas *seguramente* não obriga ao fato e à gravata. (2010, http://avesso-do-avesso.blogspot.com/2010_04_01_archive.html)
'There must be a dress code, and I find that acceptable, but it surely does not force one to wear a suit and a tie.'

Since uncertainty markers are incompatible with the strengthening use, we must conclude that Portuguese *seguramente* is either a marker of certainty or an ambivalent epistemic expression. The latter seems to be the case: according to Hummel (2018) the cognates of *seguramente* in the main Romance languages have developed uncertainty semantics. In order to conclusively prove the claim that *seguramente* and equivalents have become pure markers of probability only in the languages of Spain, it is crucial to know whether the adverb in other Romance languages, such as French and Italian, still features the strengthening use. As far as I know, no relevant data is available.

If we assume that the emergence of *seguramente* as a pure marker of uncertainty is a phenomenon of the languages of Spain, does this mean that it is an innovation of Spanish that spread over its most immediate area of influence? Of course, Spanish plays a leading role in the exchange of linguistic elements through language contact in the context of the Spanish State (see Gugenberger et al. 2013 for an overview of the Galician situation). Furthermore, one cannot ignore the facts that (i) *-ment(e)* adverbs were learned, (mostly) written forms until the 20th century and (ii) many of the languages in (19) were scarcely written before that time. Nevertheless, a "loanword" scenario would be an oversimplification (at least for the Middle Ages and from the late 19th century onward), and a more realistic picture is one where speakers/writers, taking part in the same discourse tradition, make different languages interact with each other through their own linguistic activity. In fact, it would be inadequate to talk about linguistic borrowing, since we are confronted with the reactivation of (pre)existing linguistic devices. This reactivation is a case of "linguistic convergence," which is triggered by "communicative convergence," a situation where "contact speakers are saying something

new, communicating an idea that was hitherto not usually expressed by speakers of that language" (Otheguy 1995:219). In such a scenario, one would expect parallel developments between languages if contact remains stable over time. This is, I contend, the case of *segurament(e)*.

The Galician data above show patterns not reported for Spanish. It is unlikely that this is a consequence of different evolutions. Rather, it may be due to different research interests and analytical criteria in the works dealing with the matter. According to Villar Díaz (2013) (see also Suárez Hernández 2018), the current epistemic value of Spanish *seguramente* can be detected in ambiguous instances between the 16th and 18th centuries, and spread during the 19th century. This author neglects the difference between epistemic, strategic, and pragmatic uses, and puts an end to her analysis with the emergence of the probability meaning. The examination of Galician data conducted in this paper adds complexity to the contemporary situation, showing that the adverb was an ambivalent epistemic form between the 1880s and the 1920s. If the same criteria were to be applied to Spanish, similar results would probably follow.

Likewise, the Spanish data reveal the missing pieces in the Galician puzzle. After the initial medieval period in which *seguramente* was used to convey certainty, it became epistemically weaker, conveying both certainty and uncertainty. We can say that this situation lasted several centuries, probably from the 16th to the beginning of the 20th century. This explains why strategic and pragmatic uses are so scarce in the data examined: the period from the 1880s to the 1920s represents the final moments of the transition from a weakened certainty expression to an uncertainty marker. In accordance with this is the fact that the reference dictionary for Spanish includes 'probably' as a meaning for *seguramente* for the first time in its 1927 edition (Polo 2014). Prescriptive dictionaries are not known for going along with linguistic change, so the probability meaning would presumably have been too prominent in the 1920s to be ignored. In fact, no clear certainty epistemic use was found in the 20th century, which may mean that the conventionalization of probability was already far advanced, and that the marginal strengthening and pragmatic uses of this period were outdated.

5 Final Remarks

The semantic evolution of *seguramente* consists of two main phases. In the first one, the original manner use led to a strengthening function, which in turn set the path for an epistemic modal meaning. In the second phase, the adverb underwent a semantic change involving the weakening of the epistemic meaning, that is, leading to a lower value on the epistemic scale. With the exception of the shift from manner to discourse marker, these changes are either alien or opposite to the well-established tendency in semantic change known as (inter)subjectification. This fact may be taken to mean that (inter)subjectification can be circumvented if more powerful factors are at play, namely: the way linguistic items are used in discourse (e.g., in the case at issue, to reinforce something the speaker is not sure about), and the paradigmatic pressure to avoid synonymy.

In explaining the historical path from certainty to uncertainty, several types of markers have been identified, which are the result of different combinations of epistemic values and discourse functions. Items coding epistemic certainty can be used as strengthening devices; those coding uncertainty can be used as mitigators; and epistemically ambivalent markers (coding certainty or uncertainty depending on the context) can be used to yield a special kind of interrogative. The (apparently contradictory) character of the latter type is the natural result of the accumulation of functions caused by semantic change.

Shared discourse traditions between European languages in general and between the languages of Spain in particular help us to better understand the history of *-ment(e)* adverbs. It seems that epistemic and discourse markers easily spread across languages through shared discourse traditions. In fact, items from different languages appear to evolve simultaneously in the same direction, which points towards discourse traditions, rather than national languages, as the macro-locus of language change.

Further research includes diachronic studies on *segurament(e)* in the Spanish context, and synchronic studies on the contrast between present-day *seguramente* and *probablemente*. It would also be interesting to look at *sure(ly)* and equivalent lexical bases in other languages, since it is unclear whether their weakening when used as epistemic items is a general cross-linguistic trend or an European phenomenon.

Acknowledgments This work has received financial support from the Xunta de Galicia and the European Union (European Social Fund – ESF). Thanks are due to the editor, an anonymous reviewer, and Francisco Cidrás for insightful comments.

References

Academia de la Llingua Asturiana (2015). *Diccionariu de la Llingua Asturiana*. Uviéu: Academia de la Llingua Asturiana.

URL http://academiadelallingua.com/diccionariu/

- Aijmer, K. (2002). Modal adverbs of certainty and uncertainty in an English-Swedish perspective. In H. Hasselgård, S. Johansson, B. Behrens, & C. Fabricius-Hansen (Eds.) *Information Structure in a Cross-linguistic Perspective*, (pp. 97–112). Amsterdam / New York: Rodopi.
- Álvarez, R., & Xove, X. (2002). Gramática da lingua galega. Vigo: Galaxia.
- Brown, P., & Levinson, S. C. (1987). *Politeness: Some Universals in Language Usage*. Cambridge: Cambridge University Press.
- Byloo, P., Kastein, R., & Nuyts, J. (2007). On certainly and zeker. In M. Hannay, & G. J. Steen (Eds.) Structural-Functional Studies in English Grammar: In Honour of Lachlan Mackenzie, (pp. 35–57). Amsterdam / Philadelphia: John Benjamins.
- Caffi, C. (1999). On mitigation. Journal of Pragmatics, 31(7), 881-909.
- Carretero, M. (2012). Los adverbios ingleses *certainly, naturally* y *surely,* y su traducción al español. In P. Martino Alba, & C. Lebsanft (Eds.) *Telar de traducción especializada,* (pp. 141–152). Madrid: Dykinson.
- Company Company, C. (2017). Adverbial adjectives and *-mente* adverbs face to face. In M. Hummel (Ed.) *Adjective Adverb Interfaces in Romance*,

(pp. 257–286). Amsterdam / Philadelphia: John Benjamins.

- Croft, W. (2000). *Explaining Language Change: An Evolutionary Approach*. Harlow: Longman.
- Dahl, Ö. (2000). The grammar of future time reference in European languages. In Ö. Dahl (Ed.) *Tense and Aspect in the Languages of Europe*, (pp. 309–328). Berlin / New York: Mouton de Gruyter.
- Dicionário Priberam da Língua Portuguesa (2013).

```
URL https://priberam.pt/dlpo/
```

- Downing, A. (2001). "Surely you knew!": *Surely* as a marker of evidentiality and stance. *Functions of Language*, 8(2), 251–282.
- Downing, A. (2008). From transparency to opacity: *Surely*, intersubjectivity and English cultural norms. In M. J. Lorenzo Modia (Ed.) *Proceedings 31st AEDEAN Conference*. A Coruña: Universidade da Coruña.
- Euskaltzaindia (2016). Euskaltzaindiaren Hiztegia. Bilbo: Euskaltzaindia.

URL http://www.euskaltzaindia.eus/

- Ferreiro, M. (2001). *Gramática histórica galega*, vol. 2. Santiago de Compostela: Laiovento, 2 ed.
- Freixeiro Mato, X. R. (2003). *Gramática da lingua galega*, vol. 4. Vigo: A Nosa Terra.
- Gugenberger, E., Monteagudo, H., & Rei-Doval, G. (Eds.) (2013). *Contacto de linguas, hibrididade, cambio: contextos, procesos e consecuencias*. Santiago de Compostela: Consello da Cultura Galega / Instituto da Lingua Galega.
- Halliday, M. A. K. (2004). *An introduction to Functional Grammar*. London: Hodder Arnold, 3 ed.
- Hummel, M. (2013). La dimensión intercultural de la expansión diacrónica de los adverbios en *-mente*. In M. P. Garcés Gómez (Ed.) *Los adverbios con función discursiva: Procesos de formación y evolución*, (pp. 15–41). Madrid / Frankfurt am Main: Iberoamericana / Vervuert.
- Hummel, M. (2018). Romance sentence adverbs in *-mente*: Epistemic mitigation in synchrony and diachrony. *Linguistik Online*, *92*(5), *111–144*.
- Institut d'Estudis Catalans (2007). *Diccionari de la llengua catalana*. Barcelona: Institut d'Estudis Catalans.

URL https://dlc.iec.cat/

- Lenker, U. (2010). Argument and Rhetoric. Adverbial Connectors in the History of English. Berlin / New York: De Gruyter Mouton.
- López-Couso, M. J. (2010). Subjectification and intersubjectification. In

A. H. Jucker, & I. Taavitsainen (Eds.) *Historical Pragmatics*. Berlin / New York: De Gruyter Mouton.

Mariño Paz, R. (2008). Historia de la lengua gallega. Munich: Lincom.

- Nuyts, J. (2001). *Epistemic Modality, Language, and Conceptualization*. Amsterdam / Philadelphia: John Benjamins.
- Nuyts, J. (2005). The modal confusion: On terminology and the concepts behind it. In A. Klinge, & H. Høeg Müller (Eds.) *Modality: Studies in Form and Function*, (pp. 5–38). London / Oakville, CT: Equinox.
- Nuyts, J. (2014). Notions of (inter)subjectivity. In L. Brems, L. Ghesquière, &
 F. Van de Velde (Eds.) *Intersubjectivity and Intersubjectification in Grammar* and Discourse: Theoretical and Descriptive Advances, (pp. 53–76). Amsterdam / Philadelphia: John Benjamins.
- Otheguy, R. (1995). When contact speakers talk, linguistic theory listens. In E. Contini-Morava, B. S. Goldberg, & R. S. Kirsner (Eds.) *Meaning as Explanation: Advances in Linguistic Sign Theory*, (pp. 213–242). Berlin / New York: De Gruyter.
- Polo, A. (2014). Adverbios epistémicos en las dos primeras traducciones al castellano de *Dei delitti e delle pene*. In M. C. Grande López, L. Martín Aizpuru, & S. Salicio Bravo (Eds.) *Con una letra joven: Avances en el estudio de la historiografía e historia de la lengua española*, (pp. 99–104). Salamanca: Ediciones Universidad de Salamanca.

Real Academia Española (2014). *Diccionario de la lengua española*. Madrid: Real Academia Española, 23 ed.

URL http://dle.rae.es/

Real Academia Galega (2012). *Dicionario da Real Academia Galega*. A Coruña: Real Academia Galega.

URL http://academia.gal/dicionario

- Rivas, J., & Sánchez-Ayala, I. (2012). Procesos de gramaticalización en el desarrollo de las aportaciones reactivas: el caso de *efectivamente*. In E. Montero Cartelle, & C. Manzano Rovira (Eds.) *Actas del VIII Congreso Internacional de Historia de la Lengua Española*, vol. 1, (pp. 2363–2374). Santiago de Compostela: Meubook.
- Santamarina, A. (2014). *Tesouro informatizado da lingua galega*. Santiago de Compostela: Instituto da Lingua Galega.

URL http://ilg.usc.es/TILG/

Simon-Vandenbergen, A.-M. (2007). No doubt and related expressions: A

functional account. In M. Hannay, & G. J. Steen (Eds.) *Structural-Functional Studies in English Grammar: In Honour of Lachlan Mackenzie*, (pp. 9–34). Amsterdam / Philadelphia: John Benjamins.

- Simon-Vandenbergen, A.-M., & Aijmer, K. (2007). *The Semantic Field of Modal Certainty: A Corpus-based Study of English Adverbs*. Berlin / New York: Mouton de Gruyter.
- Suárez Hernández, A. (2018). *Análisis diacrónico de adverbios con función discursiva: Hacia una descripción lexicográfica*. San Millán de la Cogolla: Cilengua.
- Traugott, E. C. (2010). (inter)subjectivity and (inter)subjectification: A reassessment. In K. Davidse, L. Vandelanotte, & H. Cuyckens (Eds.) Subjectification, Intersubjectification and Grammaticalization, (pp. 29–71). Berlin / New York: De Gruyter Mouton.
- Traugott, E. C. (2014). On the function of the epistemic adverbs *surely* and *no doubt* at the left and right periphery of the clause. In K. Beeching, & U. Detges (Eds.) *Discourse Functions at the Left and Right Periphery: Crosslinguistic Investigations of Language Use and Language Change*, (pp. 72–91). Leiden / Boston: Brill.
- Traugott, E. C., & Dasher, R. B. (2002). *Regularity in Semantic Change*. Cambridge: Cambridge University Press.
- Varela Barreiro, X. (2007). Tesouro medieval informatizado da lingua galega. Santiago de Compostela: Instituto da Lingua Galega. URL http://ilg.usc.es/tmilg
- Vázquez Rozas, V. (2010). Sobre *seguro*: Del riesgo presupuesto a la certeza epistémica. In M. J. Rodríguez Espiñeira (Coord.) (Ed.) *Adjetivos en discurso. Emociones, certezas, posibilidades y evidencias,* (pp. 109–146). Santiago de Compostela: Universidade de Santiago de Compostela.
- Villar Díaz, M. B. (2013). La evolución de los adverbios y locuciones adverbiales de modalidad epistémica. In M. P. Garcés Gómez (Ed.) Los adverbios con función discursiva: Procesos de formación y evolución, (pp. 158–199). Madrid / Frankfurt am Main: Iberoamericana / Vervuert.

The Semantics of Possessive Noun Phrases and Temporal Modifiers

Emanuel Quadros

Abstract This paper examines the semantics of possessive NPs like *Joana's former mansion*. I argue that the scope ambiguity in the interpretation of temporal modifiers like *former* supports a two-place approach to the semantics of possessive noun phrases, according to which they are always formed from relational denotations. This ensures that a possessive relation is available in the syntactic/semantic composition at the point where a temporal modifier is added. In this way, the modifier can take scope over the relation. I derive the difference between *ex-* and syntactic modifiers, with respect to the kinds of possessive relations they can modify, from the distinction between sortal and relational nouns. The prediction is that derivational affixes can only target relations that are lexically encoded, not those arising via type-shifting.

Keywords possessive \cdot noun phrase \cdot type-shifting \cdot morphology \cdot semantics \cdot temporal modifier \cdot Combinatory Categorial Grammar

```
E. Quadros, Yale University, http://campuspress.yale.edu/manuquadros/
```

In Christopher Pinon (Ed.), Empirical Issues in Syntax and Semantics 12, 247–276. Paris: CSSP. http://www.cssp.cnrs.fr/eiss12/ © 2019 Emanuel Quadros

1 Introduction

This paper is concerned with the semantics of possessive NPs and its interaction with temporal modifiers like *former* and *ex*-. I will focus specifically on the so-called Saxon genitive, exemplified in (1). Despite the label "possessive," an indefinite number of relations between entities can be described by possessive constructions, such as ownership, control, kinship and part-whole relations, to name a few. We find some freedom in the range of possible relations even in the case of phrases like (1), despite the overwhelming preference to interpret it as referring to the individual who stands in the 'fatherhood' relation to John. For example, in the context of a coaching program for fathers, where John works as a coach, an utterance of (2) would be acceptable.

- (1) John's father
- (2) John's father has made a lot of progress.

As is usual in the literature, I will refer to the relations that can be instantiated in a possessive noun phrase as *possessive relations*, even when they do not refer to possession proper.

There is some debate in the literature as to which possessive relations are determined by the lexical items involved in a possessive phrase and which are fully derived by pragmatic means. Vikner & Jensen (2002), for instance, argue for a greater role of lexical information than most approaches, exploiting the content of structured lexical entries to derive relational denotations. On the other extreme, Peters & Westerståhl (2013) argue that all possessive relations are pragmatically derived, and are present in the semantic composition only as a free variable. On the basis of restrictions on the availability of possessive relations in the context of different modifiers, I argue for an account that takes insights from both of these positions: one in which most relations are contextually derived, but some have to be specified in the lexicon. A large part of the empirical basis for my account comes from restrictions on possessive relations under the scope of temporal modifiers, exploring in particular how and why morphological and syntactic modifiers differ in the range of relations with which they are compatible.

Another question tackled by this paper is the scopal properties of temporal modifiers, which can take scope over the noun to which they attach or over the relation between the possessed entity and its possessor. For example, the noun phrase in (3b) can refer to a house where I used to live, and this is probably the most salient reading of this phrase for speakers who find it acceptable at all. (3b) can also refer to an entity that was formerly a house, but may still be
mine in some sense. The second interpretation may not be readily available for all speakers, and it is certainly more accessible in the case of (3a). The important point here is that both *ex-* and *former* are able to modify the relation between the head noun and the possessor, even though both modifiers are attached (morphologically or syntactically) to the head noun only.

(3) a. My former house b. My ex-house

One motivation for paying attention to these facts, at least in the case of *ex-*, is that this sort of scope interaction poses a challenge to some versions of the Lexical Integrity Hypothesis, according to which word-level morphological elements should not directly interact with syntactic elements (Lieber & Scalise 2006). As the discussion of (3b) above shows, the prefix may have phrasal scope over the possessive pronoun, even though lexical integrity would imply that sub-lexical elements cannot enter into scopal relations with elements outside of their lexical domain. As we will see later in this paper, the strength of this counterexample depends on the semantic analysis of these constructions. More specifically, I will argue that this problem dissolves once we analyze possessive relations as being available in the denotation of the noun, instead of being introduced by the possessor or by the possessive morpheme.

From the perspective of non-transformational theories of the interface between semantics and morphosyntax, additional interest in this topic comes from the need to derive this sort of scope ambiguity without resorting to syntactic movement. Although the substance of my analysis is compatible with other syntactic frameworks, I couch it in a version of Combinatory Categorial Grammar (CCG). I will argue that the flexible constituency available in CCG allows us to directly derive coordination facts that are challenging for more rigid phrase-structure approaches.

In section 2, I review a number of approaches to the semantics of possessive noun phrases. The main point of contention in this section

will be the mechanism by which possessive relations are introduced in the semantic composition. I will argue that the interaction of temporal modifiers with possessive relations favors a type-shifting approach in which sortal nouns acquire a relational denotation in the context of a possessive morpheme, as proposed by Vikner & Jensen (2002), among others.

In section 3 I discuss the extent to which possessive relations are free, reviewing restrictions proposed in early work by Barker (1995), and presenting novel findings in this area. This topic is also taken up in section 4 in connection with the range of relations that can appear under the scope of different types of temporal modifiers. Finally, in section 5, I provide a grammar fragment integrating all the aspects of my analysis.

2 Approaches to the Semantics of Possessive NPs

As discussed in Löbner 1985, common nouns can have two basic interpretations, sortal or relational. Sortal interpretations are those usually represented by a one-place function, conventionally denoting a set of entities. This is the interpretation given, for example, to the noun table in most contexts. Relational interpretations, in contrast, denote relations between entities, and can be represented by two-place functions. Nouns are often compatible with both kinds of interpretation, so, even though *table* is prototypically a sortal noun, the phrase *my table* denotes an object that stands in some relation to me. In this case, we can say that *table* has a relational use by virtue of being in a possessive construction. Other nouns, however, seem to encode a relation intrinsically, such as sister, husband, edge and boyfriend. These nouns can be said to be prototypically relational, in the sense that an entity can only be described by one of them if it stands in the relevant relation to some other entity; for example, someone can only be described by the noun *boyfriend* when they are in a certain conventional relation to some other person.

Possessive constructions can be formed either from prototypically sortal nouns, of the *table* kind, or from relational nouns of the *boyfriend* kind. In trying to give a unified account of the facts, approaches to the semantics of possessives have differed on which of these cases is the basic one. One-place approaches assume that possessor phrases always combine with nominal predicates denoting a set, of type $\langle e, t \rangle$ (hence "one-place"). Therefore, approaches of this kind then need some mechanism to allow the formation of possessive noun phrases from relational nouns, under the reasonable assumption that these have a lexical denotation of type $\langle e, \langle e, t \rangle \rangle$. Two-place approaches, on the other hand, assume that possessor phrases always combine with nominal predicates denoting a relation, of type $\langle e, \langle e, t \rangle \rangle$. The converse is true in this case: two-place approaches need some mechanism allowing for the formation of possessive NPs from nouns that are not lexically relational.

I argue for a two-place approach for English possessive noun phrases in this paper, mainly on the basis of their interaction with temporal modifiers. In the remainder of this section, I review and reject recent arguments for a one-place approach and sketch the basic elements of my analysis.

2.1 Arguments for a One-Place Approach

The most recent example of a one-place approach to the semantics of possessives is the work of Peters & Westerståhl (2013). In their proposal, all nouns denote a set at the N'-level, including relational ones. When a noun is related to a possessor, the possessive relation is in all cases introduced by the possessive morpheme. Since the content of the relation is contextually determined, what the possessive morpheme introduces is actually a relational variable. When the head noun is a relational one, the value taken by this variable may or may not coincide with the relation prototypically associated with the head noun.

The authors advance two arguments for this position. The first one is that relational nouns have the same distribution as other nouns. This point was also raised by Partee (2011) as a problem for the categorization of relational nouns as transitive common nouns. She illustrates the problem with conjoinability tests like the one in (4). Under the usual assumptions about conjunction, this sentence should not be possible if relational and non-relational nouns have different syntactic categories and semantic types.

(4) John has a good job, a nice house, a beautiful wife, clever children, and plenty of money (and an ulcer). (Partee 2011:546)

The second argument given by Peters & Westerståhl for choosing a one-place approach is related to the freedom of possessive relations. Even in the case of relational nouns, the authors note, context must be invoked in order to find the appropriate relation for the interpretation of a possessive noun phrase. The appropriate relation may be the one predicted by the lexical specification of the noun, but it may also be another one. In Barker's (2011) example, *John's brother* may be a brother that was somehow assigned to John (for example, suppose that John is one of a group of journalists assigned to profile each of the brothers of a famous person). These cases show that some operation for deriving a set from a relation is necessary in any case, such as "projection," as suggested by Peters & Westerståhl, by means of which a set of brothers can be obtained from the domain of the brother relation.

As the argument goes, since context can always override the lexical preferences of relational nouns, and since a mechanism of projection must be available in any case, the advantage of having a possessor phrase combine with relational denotations to derive lexical interpretations directly would be illusory. Hence, in favor of uniformity, the authors chose an analysis in which possessor phrases always combine with a set, and in which the relation between possessor and possessee is introduced by the possessive morpheme and determined by context.

It is important to note, however, that Peters & Westerståhl do not argue against the existence of lexically relational nouns, because there are good examples of nouns taking complements outside of the possessive domain. Since such nouns can also head possessive noun phrases, their approach has to assume a semantic projection operation that creates sortal denotations from lexically relational nouns. It is only then that a noun of this kind, with a derived sortal denotation, can combine with a possessor phrase, which will introduce a relation between the possessor and the possessee.

The fact that, in this approach, relational nouns lose their relationality only to have it restored when they combine with a possessor, mirrors a feature of two-place approaches that Peters & Westerståhl criticize. In two-place approaches such as the one argued for in this paper, relational nouns combine directly with possessor phrases, so that the resulting possessive relation comes from the lexical meaning of the possessee. Then, the issue arises of how to account for cases in which a relational noun is interpreted with a relation other than that predicted by its lexical meaning. The approach taken in Barker 2011 involves a mechanism of semantic projection not unlike the one assumed by Peters & Westerståhl; namely a detransitivizing typeshifting operation that results in a set corresponding to the entities in the domain of the relevant relation. The resulting sortal denotation can then be shifted back into a relational denotation, now a contextually controlled one, in order to be compatible with a possessor.

On the face of it, the latter move seems suspicious, since two typeshifting operations going in opposite directions apply to the same lexical item; first turning a relational denotation into a sortal one, then back into a relational one. Phonologists will recognize the similarity to a "Duke of York gambit" (the kind of analysis where a derivation has the general form $A \rightarrow B \rightarrow A$). But as Pullum (1976) argued, a general prejudice against this sort of derivation is unfounded, since the doing and undoing of the intermediate step is often motivated.

Moreover, the shift from relational nouns to sortal, then back to relational, is also found in Peters & Westerståhl's approach, as I pointed out earlier. Hence, there is no argument from parsimony here. The difference is that, in their one-place approach, the last step in the derivation, which reintroduces relationality, is effected by the possessive morpheme. Strictly speaking, the possessee does not become relational again in their approach, but since it is put into a relational configuration, the derivation and its effects are very similar when we only consider non-modified possessives like *John's brother*. The way these alternatives can be distinguished is by examining the empirical consequences of assuming that possessees acquire a relational denotation only after combining with a possessive morpheme, as compared to assuming that relationality is already present in the semantic composition at that point. In section 2.2, I will argue that the second alternative gives a more directly compositional account of the interaction of possessive relations with temporal modifiers, and should therefore be preferred.

2.2 Arguments for a Two-Place Approach

Peters & Westerståhl (2013) discuss and reject one argument given by Partee & Borschev (2003) against a one-place approach, based on the semantics of *former*. Consider the possessive NP in (5). A oneplace account will not readily get the wide-scope reading of *former*, in which *Mary's former mansion* refers to something that is still a mansion, but is not owned by Mary anymore. The reason is that the possessive relation between these two entities would not be available in the semantic composition at the point at which the temporal modifier combines with the possessed noun, since the relation is introduced by the possessive morpheme.¹

(5) Mary's former mansion was destroyed by fire.

In a two-place approach, possessor phrases combine with relational denotations, saturating an argument role already present in the denotation of the possessed noun. The question for two-place

¹This problem also applies to mixed approaches in which possessors can combine both with relational and with sortal nouns. Note that (5) exemplifies the case of a sortal noun. In a mixed approach, just as in the one-place approach of Peters & Westerståhl (2013), *mansion* would be a one-place predicate throughout the derivation; relationality would be introduced in the construction by the possessive morpheme, hence outside the scope of *former*.

approaches, then, is how a possessive relation can be introduced in the case of nouns that are lexically sortal. The route taken in much of the literature, including this paper, is the postulation of a typeshifting operation that turns one-place nominal predicates into twoplace relations. We will explore how this works later in the paper, but for now, a crucial benefit of this assumption is that a possessive relation can be already present in the semantic composition by the time *former* combines with the possessed noun; hence the temporal adjective can scope over this relation.

Peters & Westerståhl recognize that a one-place account will not get the wide-scope reading of *former* in (5), but they argue that a twoplace account would also not give the right result in this case. The reason is that applying *former* to a relational denotation such as (6a), derived via type-shifting, would result in a representation like (6b), under the assumption that an operation applying to a conjunction commonly applies to both conjuncts.

- (6) a. mansion(x) \land own(y, x)
 - b. former(mansion(x)) \land formerly(own(y, x))

Partee & Borschev (2003:95) do state that *former* could "in principle target either part [of the conjunction in the denotation of a shifted noun; EQ], depending on what was presupposed and what was focussed in the given context." They provide representations similar to those in (7). Even though this analysis has the technical problem of assuming, and having to ensure, that *former* targets only one part of the conjunction, I believe the spirit of the approach is correct, in that the semantic operation performed by *former* can be relevant to one conjunct or the other, or both, depending on what is relevant in a context. However, we can arrive at this result by applying *former* to the whole conjunction once we explore in more detail the semantic effect of this modifier.

- (7) a. $(PAST(mansion(x)) \land possessed-by(y, x))$
 - b. $(mansion(x) \land PAST(possessed-by(y, x)))$

A first approximation of the truth conditions of Mary's former mansion, under the assumption that the modifier targets both conjuncts in the logical form, is given in (8). As is clear from (8), we take an application of former in Mary's former mansion to describe a state of affairs in which the entity described as a mansion stood in a certain relation to Mary at some point t' prior to the reference time t. Furthermore, this state of affairs is described as not holding anymore at t. But since the relevant state of affairs is a conjunction of two subformulas, there are three ways in which it could be said not to hold at the reference time *t*: it may be the case that the entity is not a mansion; it may not stand in relation to Mary; or it may be the case that it is not a mansion and also does not stand in the specified relation to Mary. If these readings are available, reflecting the different ways a conjunction can be false, then two-place approaches, in which the possessive relation can be directly modified by former, give an insightful account of the truth-conditions of the facts. Consequently, Peters & Westerståhl's criticism dissolves.

(8) former(mns(x) \land own(m, x)), expanded as: \neg (mns_t(x) \land own_t(m, x)) \land \exists $t'(t' \prec t \land$ mns_{t'}(x) \land own_{t'}(m, x)) (Where t is some reference time)

Let us examine each of these cases with reference to the sentence (5), repeated below in (9). For convenience, let M be the formula $(mansion(x) \land own(m, x))$, where m refers to Mary and x refers to some entity that can be described as a mansion and may be owned by Mary. According to (8), M must be false if *former* is to be used as a modifier in *Mary's former mansion*. (10a) describes a case in which M is false because its first conjunct is falsified: the entity owned by Mary is not a mansion anymore. (10b) describes a case in which the second conjunct of M is false, since the mansion (which is still a mansion at least up to destruction) is not owned by Mary anymore. Finally, (10c) describes a case in which none of the conjuncts of M hold, since the property that was destroyed is not a mansion and is not Mary's anymore. All of these readings are compatible with

the analysis put forward in this paper, hence we see that when we properly define the semantic contribution of *former*, Peters & Westerståhl's semantic argument against Partee & Borschev's two-place approach disappears.

- (9) Mary's former mansion was destroyed by fire.
- (10) a. Mary used to own a mansion, which she turned into a bed and breakfast. She still owns the property, but it was recently destroyed.
 - b. Mary used to own a mansion, which she sold. The mansion was recently destroyed.
 - c. Mary used to own a mansion, which she turned into a bed and breakfast and then sold. The property was recently destroyed.

We still have to comment on the syntactic issue that Peters & Westerståhl take to be an argument against two-place approaches, namely why relational and sortal nouns have similar distributions although they are assigned to distinct types in the lexicon. As we will see in the next section, relational nouns are compatible with postnominal possessors whereas sortal nouns are not, so their distributions are not strictly the same. Nevertheless, we can freely coordinate nouns from these two classes, as we saw in (4). This is not an issue for the approach taken in this paper due to the availability of type-shifting operations taking sortal nouns to relational denotations, and relational nouns to sortal denotations. These operations can apply to resolve type-mismatches arising in the coordination of nouns from these two classes.

2.2.1 Postnominal Possessors

The flexible typing of nouns in this approach faces a small problem when we consider possessive NPs whose possessor is introduced by an *of*-phrase, as in (11a) and (11b). The examples in (11) show that *of*-phrases have to be compatible with the type of relational nouns like *friend*. But since all common nouns potentially have a re-

lational denotation in our approach, we would expect *of*-phrases to also be compatible with any common noun. This result is incorrect, as shown by (11c) and (11d). In general, extrinsic possessive relations cannot be expressed by *of*-phrases in English, except when the possessor is also marked by *'s*, as in (11e). Because of this restriction, Barker (1995:9) uses the availability of a possessive *of*-phrase as a diagnostic of whether a noun is lexically relational or not. Hence, (11) would show that *keyboard* and *fire truck* are not lexically relational.

- (11) a. a friend of Joana
 - b. a child of Joana
 - c. *a keyboard of Joana
 - d. *a fire truck of John
 - e. a keyboard of Joana's, a fire truck of Joana's

To explain this restriction, we assume that the type-shifting operation deriving relational denotations for common nouns does not apply freely. Relational denotations can be lexically specified or arise via coercion in the context of a possessive morpheme. This is the crucial assumption preventing common nouns to combine with possessive of-phrases. In the absence of a possessive morpheme, common nouns do not have the correct syntactic category, nor the semantic type, to take an *of*-phrase as a complement. This is shown in (12), where the only possible combination would be one in which the ofphrase is a modifier of keyboard. Examples in which an of-phrase is added as a modifier are given in (13). It is not always easy to identify whether a construction with an of-phrase is possessive or not. A reasonably good test in this case is the availability of a prenominal possessive. The modifiers in (13) do not have a prenominal counterpart, in contrast to the of-possessives in (11a-11b), which could be expressed by a prenominal possessor phrase.²

²The weakness of this test is that the availability of a prenominal possessive does not entail that a corresponding *of*-phrase is a complement. Recall that, in our approach, any common noun, including *keyboard*, can receive a relational denotation under coercion from a possessor phrase marked with 's. So, in principle,

- (12) keyboard: N of Joana: $N \setminus N$
- (13) a. "I found the keyboard of my dreams. Put it in this netbook and this would be the computer of my extatic bliss..." (https://ploum.net/220-board-of-columns-ofkeys/, accessed on 2017-03-23)
 - b. "The above picture is a piano keyboard of 88 keys, containing 7 1/3 octaves." (http://harmoniumnet.nl/klavier-keyboard-E NG.html, accessed on 2017-03-23)

In contrast to common nouns, lexically relational nouns subcategorize for expressions of the syntactic category of of-phrases. Thus, in (14), friend can take of Joana as its complement, giving the correct result. Note that the complement in (14) has the same category we ascribe to noun modifiers in CCG. Dowty (2003) calls complements of this type subcategorized adjuncts. The reason is that an expression of category $N \setminus N$ can function as a true modifier in some configurations, but can be subcategorized for (taken as a complement) in others. For Dowty, the categorial identity between subcategorized adjuncts and true modifiers is not an accident, for it is frequently the case that the same kinds of expressions can appear in both functions. The author provides a list of examples of expressions that can correspond to adjuncts or complements in English, of which we single out the following. In (15a), the *with*-phrase is clearly an adjunct, but is arguably a complement in (15b). The same dual function of of-phrases can be seen in the case of modifiers and complements of

a noun phrase introduced by a prepositional phrase modifying a common noun, as in (13a), could appear as the prenominal possessor of a transitivized common noun, as in *?my dreams's keyboard*. It is unclear if what makes this phrase anomalous is the lack of a reading for it or the competition with, and the preference for, the postnominal version – in any case, I expect that some speakers should find this example acceptable. The test still works, in many cases, because lexically relational nouns are guaranteed to be compatible with a prenominal possessor whenever there is a corresponding *of*-phrase introducing the same possessor as a complement of the relational noun.

nouns in (16a) and (16b), respectively.

- (14) *friend*: $N/(N \setminus N)$ of Joana: $N \setminus N$
- (15) a. John swept the floor *with a broom*b. John loaded the truck *with hay*
- (16) a. This is a piano keyboard of 88 keys.
 - b. I am the owner of 88 keys.

3 The Limits of Freedom

Freedom of the possessive relation is a characteristic property of possessive constructions in English. Regardless of our choice between ways of combining possessors and possessees, we have to account for the way in which possessive relations are integrated in the semantic composition. Since the possessive relation is not always determined by the lexical properties of the nouns involved, most authors represent it as a free parameter in the semantics of possessive constructions. The setting of this parameter involves pragmatic reasoning, although the precise mechanism is much less explored in the literature, and this paper is no exception to that.

Despite the fact that possessive relations are generally free, Barker (1995) noticed some interesting asymmetries on their expression. For instance, part-whole relations are not easily reversible, as shown in the examples below. The contrast is, to some extent, predicted by Barker's approach, since *leg* and *cover* are relational nouns, whereas *table* and *box* are not. However, as Barker notes, it is still mysterious why the relational parameter of the possessive construction, which is necessary to account for extrinsic interpretations, cannot take on a part-whole relation in these cases. Translated to our approach, the puzzle is why in (17b) and (18b), *table* and *box* cannot shift into relational denotations and have their relational variables set to the inverse of the relations we find in (17a) and (18a), respectively.

(17) a. the table's legb. #the leg's table

- (18) a. the box's cover
 - b. #the cover's box

Barker suggests that relational variables can never take on the value of a lexical possessive relation, understood as one that is encoded in a relational noun present in the construction. In (17b), for example, *table* would not be able to shift into a relational denotation whose relational variable is a part-whole relation, and have its part argument filled by *leg*, because there already is a noun lexicalizing the relevant part-whole relation, namely *leg*. For a parallel reason, *box* would not be able to shift into a relational denotation in (18b).

We will see below that this is not an absolute restriction, but first let us consider a deeper question posed by this asymmetry. Namely, why is there, to begin with, an asymmetry with respect to which relatum is lexicalized as the relational noun? I believe the explanation for this fact goes along the lines of the Nominal Argument Selection Principle of Barker & Dowty (1993:55–56), which states that "the argument for which the predicate denoted by the noun entails the greatest number of Proto-Whole properties will be lexicalized as the object of the preposition *of* or as the prenominal possessor; the argument having the greatest number of Proto-Part entailments will be lexicalized as the head argument." Proto-Whole and Proto-Part here are understood as proto-roles that are responsible for nominal argument selection, and are defined by the entailments in (19) and (20).

- (19) Proto-Part entailments
 - a. located at or defines a boundary of the other relatum
 - b. is a property of the other relatum
- (20) Proto-Whole entailments
 - a. entirely contains the other relatum as a proper part
 - b. is a concrete entity

In the case of *box* and *cover*, for instance, it is clear that when these

two nouns are related at all, *cover* describes something that is located at or defines a boundary of a box. Barker & Dowty (1993) predict, correctly, that *cover* can be lexicalized as relational noun, whereas *box* is unlikely to be, at least with respect to the sort of relations that obtain between boxes and covers.

This explains the differential lexical properties of pairs like box/ cover, but still leaves open the puzzle formulated by Barker (1995): since sortal nouns can generally be shifted into relational denotations, giving rise to possessive noun phrases with extrinsic relations, why is this possibility blocked in cases like (17b) and (18b)? As (21) shows, however, blocking is only partial. In (21), we have an example from a web forum, discussing a mooring cover which was delivered in a certain box. The difference in this case is that *cover* is not construed as a relational noun, not as the cover of something, but as a particularized object that might serve to cover something.

(21) Connie is the person who I have dealt with several times on parts, and she's good (in fact, **this cover's box** had "attn: Connie" written on it, so she must handle dealer parts/accessories orders too).

(http://www.keywestboatsforum.com/topic6093.ht
ml\#p48377, accessed on 2017-02-23)

Clearly, (21) is not a counterexample to Barker's generalization, since the relation between *cover* and *box* is not of the usual part-whole sort. However, examples like these point to an explanation for the puzzle of why certain relations are "reversible" in possessive noun phrases while others are not. Note that (21) was felicitously used in a context in which the cover was the familiar object of discussion. Unlike heads of possessive noun phrases, *cover* in this case had no need to be anchored by a possessor (in the sense of Prince 1981) in order to be identified in the context. Hence, it is reasonable to assume that *cover* was not used in its relational denotation in this case.

Having *cover* as the possessor in (21) is likely facilitated by the fact that that particular cover was not a cover of that particular box. The

box was just the package in which the cover was delivered. However, it is possible to invert the possessor-possessee order even when a part-whole relation is implied. Consider, for example, the following context: a carpenter is working on a number of tables, each of which is specifically designed for a different customer; by design, each set of legs only fits a particular table. Holding one of a number of unattached legs, the carpenter could direct the sentence in (22) to her assistant. To the extent that a leg can be individuated in a particular context without necessary reference to some entity in relation to which this leg stands, it is predicted that the noun referring to the leg can be used as a possessor.

(22) I am looking for this leg's table.

Asymmetries between which noun can more easily function as a possessor in a possessive NP can also be found with other types of relations, as in (23a) and (23b). The possibility of reversing the possessor-possessee order may be more or less available in each case.

- (23) a. the student's name # the name's student
 - b. this speaker's language# this language's speaker

The upshot of these considerations is that the restriction on the reversibility of part-whole relations identified by Barker (1995) is not to be found in lexical semantics, as the author suggests, but instead relates to the different discourse functions of the elements of the possessive noun phrase. An entity can serve as a possessor if it is familiar enough, in a context, to anchor an object with which it is related.³

Another restriction on possessive relations is found in their in-

³The possessee does not have to be less familiar than the possessor. In *I found someone's key*, the key, being perceptually immediate, may be more familiar to the speaker than its (maybe unknown) owner. However, the existence of an owner is implied, making them familiar enough for the possessor to be acceptable.

teraction with morphological and syntactic modifiers, and this one can be explained by the properties of the semantic composition. As suggested in the introduction to this paper, there are differences in the availability of possessive relations under the scope of different temporal modifiers. The relevant contrast is between syntactic and morphological modifiers, as exemplified in (24).

We note a sharp contrast between relations that are more clearly inherent in the meaning of the head of the construction, such as in (24a) and (24b), and those that seem to be pragmatically derived, as in the remaining cases in (24). In the case of heads like *girlfriend* and *boss*, whose meanings require the existence of some other entity of which the entity being described is a girlfriend or a boss, *ex*prefixation is perfectly acceptable. The prefix becomes much less acceptable with nouns that do not directly encode any kind of relation, as shown in (24c–24e). In Löbner's (1985) terminology, these correspond to sortal nouns and stand in contrast with relational nouns, which encode a relation.

- (24) a. My (former/ex-)girlfriend
 - b. Our (former/ex-)boss
 - c. Since we sold it, I've seen our (former/?ex-)car every single day.
 - d. The waiter moved us back to our (former/*ex-)table.
 - e. He has only the memory of his (former/*ex-)injury.
 - f. "A carved wooden peg with a brass tip replaced his (former/*ex-) leg." (*Sylvia*, Bryce Courtenay)

I advance the hypothesis that *ex*- can only modify relations that are lexically specified. However, (24f) is a problematic case for this generalization. Even though *leg* is usually taken to be a relational noun, and thus to have a relation encoded in its lexical entry, it cannot be modified by *ex*-. The same seems to be true of terms referring to other body parts. I take this to be a principled exception, suggesting that part-whole relations at the lexical level are treated as inalienable in English, and furthermore, that *ex*- is incompatible with

inalienability.

The case of (24c) also deserves comment, since *car* is lexically sortal for most speakers. For a few of the speakers I consulted, however, it is acceptable as a base for *ex*- prefixation in possessive NPs. I suggest that this variation relates to the lexical status of the distinction between relational and sortal nouns. Given that many English speakers frequently encounter *car* in possessive noun phrases interpreted as involving some form of ownership, and given that cars are normally related to an owner in our daily lives, it is likely that some English speakers have lexicalized a relational denotation for this noun, perhaps in addition to its sortal one.⁴

Despite these problematic cases, the contrast between *former* and *ex*- in (24) is clear. We see in the examples above that *former*, unlike the prefix, seems to be able to modify any kind of relation. Thus, in providing a formal analysis of the modification of possessive relations, we have to take these restrictions into account and explain why they only arise in the case of *ex*-.

4 Temporal Modification of Possessive Relations

At least since the pioneering work of Enç (1986), it is known that nominal predicates have a temporal interpretation that is not necessarily determined by the time of the verbal predication. Consider the example in (25).

(25) A: Gosh, the government is really pushing a hard line with these countries.

B: Well, the president already made it quite clear during the incident in 1980 that he wasn't a soft guy.

(Tonhauser 2002:293)

In (25), the underlined noun phrase can be interpreted at the ut-

⁴As pointed out by a reviewer, this predicts that other nouns that are conventionally related to a possessor in English-speaking cultures, like *dog* and *computer*, should show a similar behavior. While uses of *ex*- in these cases do appear in web searches, properly exploring this prediction is beyond the scope of this paper.

terance time, instead of at the verbal predication time. This follows, for Tonhauser (2002), from the fact that speaker A introduced a set of relevant individuals who are part of the government at the utterance time, and from the requirement that definite noun phrases refer to some participant already established in the context. In the absence of contextual pressures to the contrary, the verbal predication time is the default source for the temporal interpretation of noun phrases.

Besides context, another way in which the temporal interpretation of a nominal predicate can be manipulated is through the introduction of temporal modifiers, such as *former*, *future* or *present*. Following Tonhauser (2002), I take *former* to introduce a time variable with the requirement that the value of this variable be a time following the time at which the nominal predicate holds. The temporal variable introduced by the temporal modifier is itself then subject to being identified with the verbal predication time or with another contextually salient time.

(26)
$$\llbracket former \rrbracket = \lambda P \lambda t \lambda x. \neg P(t)(x) \land \exists t'[t' \prec t]. P(t')(x)$$

4.1 Morphological versus Syntactic Temporal Modifiers We observed before that *former* and *ex*- are not compatible with the same range of semantic relations, even though they have the same semantic effect of restricting the temporal interpretation of a nominal predicate to some time prior to a reference time.

I follow the spirit of Dowty's (1979) proposal on the distinction between lexical and syntactic rules, in assuming that the same set of operations is available for both kinds of rules. The difference between lexicon and syntax would be primarily one of function, not of form. While the function of lexical rules is to extend the basic set of expressions available to the grammar, syntactic rules serve to combine these basic expressions in accordance with translation rules that guarantee a compositional interpretation. In the simplest case, lexical extensions will also be fully compositional, following the translation rules provided by the grammar. But their product can deviate from full compositionality; for example, by the familiar process of lexicalization: since words can be stored in the mental lexicon, their meanings can be enriched with features that are not predicted by regular interpretation rules.

The consequence of this view for *ex*- depends on some additional assumptions. If we take this prefix to be added by a lexical rule, then it serves the role of extending the set of basic expressions of the grammar. However, extending the set of basic expressions is, in relative terms, rarely necessary, especially when there is some syntactic rule having the same effect (in this case, *former* modification). Thus, one reason why *ex*- prefixation is more restricted than *former* modification is to be found in a theory of morphological productivity.

More importantly, I hypothesize that items belonging to open lexical classes cannot be lexicalized with free variables. Hence, no sortal noun could be shifted into a relational denotation and be lexicalized as such. From this hypothesis, we derive the result that morphological elements like *ex*- cannot modify extrinsic relations, since these relations are not present in the lexicon.

To make this suggestion more concrete, let us consider the case of *Joana's table*. This possessive noun phrase can refer to a table that Joana owns, built, designed, reserved in a restaurant, is presently occupying, or to one that stands in any other plausible relation to her. Presumably, none of these relational uses of *table* is lexicalized. Instead, they arise, in the approach adopted in this paper, via a typeshifting rule that takes the denotation of the head noun as its input and returns a denotation containing a free relational variable. In (27a), we define this rule, and in (27b) we show the semantic effect of its application to the noun *table*.

(27) a. REL :=
$$N \Rightarrow N/NP : \lambda P \lambda y . \lambda x . P(x) \land \pi(x)(y)$$

b. REL(*table*) = $\lambda y \lambda x . table(x) \land \pi(x)(y)$

Note that the assumption that type-shifting operations such as (27a) can only apply to resolve type mismatches arising in the syntactic/semantic composition leads to (27a) being intrinsically ordered after *ex*- prefixation. Hence, the hypothesis that lexical items belonging to open classes cannot contain a free variable in their denotation is sufficient to account for *ex*-'s being more restricted than *former* – simply because a variable π is guaranteed not to be available at the point in the derivation in which *ex*- is attached.⁵

Further support for this suggestion comes from the domain of English compounding. As in possessive noun phrases, the relation between the components of an NN compound in English is largely free, in the sense that, for any noun combination, there is no reliable way to determine the meaning of the compound, unless one has already encountered (and interpreted) the compound before or can infer its meaning given enough contextual information. But regardless of the freedom observed in compounding as a whole, specific instances of the NN compound construction are lexicalized with a specific relation, and once one of these formations has been lexicalized for a speaker, its meaning is largely fixed (modulo meaning extensions of the sort that any lexical item is subject to). Consider, for example, steam boat, garden party, flea bite, hand brake or tear gas, which do not show the same semantic flexibility as possessive noun phrases do. While a phrase like our brake is quite open with respect to the relation obtaining between the possessor and a particular brake, in hand brake, we do not need to access any contextual information to interpret the relation between the two base nouns. If this paper is correct, this follows from the fact that NN compounds cannot be lexicalized

⁵If we had reasons to reject the assumption that type-shifting only occurs under coercion, and instead take such operations to apply freely, as suggested by Barker (2011), the restrictions on *ex*- could be derived in a similar way. Since *ex*-is introduced by a lexical rule, serving to extend the set of basic expressions, its introduction has to result in a valid member of the corresponding lexical category – in the case of interest, for categories *N* or *N*/*NP*, the result should correspond, respectively, to a set of entities (e.g., $\lambda x.fireman(x)$) or to pairs of entities taking part in a relation specified by the noun (e.g., $\lambda y \lambda x.daughter(x)(y)$). By hypothesis, a basic expression containing a free relational variable, like the one in (27b), would not be a valid lexical entry; by extension, such an expression could not be contained in the product of the lexical rule introducing *ex*-.

with a free variable corresponding to the relation between the two Ns.

Some predictions stem from this idea. We will comment briefly on them, but will not explore them further in this paper. First, unlike English, some languages show morphological marking of the operation that turns sortal nouns into relational nouns (see Aikhenvald & Dixon 2013 for an overview of the patterns). Our approach predicts that whenever it can be established that such morphological markers are derivational (as opposed to inflectional), and hence serve to extend the set of basic expressions available in the grammar, it must also be the case that they have a more specific semantics than the rule in (27a). Another prediction is that whenever we find morphological elements similar to *ex*- in other languages, we should also find that they cannot modify possessive relations that are not lexically encoded.

5 Grammar Fragment

This section presents a fragment of the grammar of possessive noun phrases in English, building on the discussion developed in the previous sections. The syntax is couched in a version of CCG (Steedman & Baldridge 2011).

The most important departure from most of the earlier treatments of possessive noun phrases is the lexical entry I propose for the possessive clitic 's, which, following Coppock & Beaver (2015), does not include any definiteness information. Also, this lexical entry is not the source of the possessive relation, which is instead part of the denotation of the possessee, as discussed in section 2. The possessive morpheme 's takes a possessor and a possessive relation as arguments, and feeds the first to the latter. As in Coppock & Beaver's treatment, the possessive morpheme has no particular semantic effect, being just an identity function operating on the possessive relation present in the denotation of the noun.⁶

⁶The fact that 's does not impose a new relation, and just transmits the relation provided by the head noun, reflects the indeterminacy of this possessive marker

(28) 's := $(NP/(N/NP)) \setminus NP : \lambda y \lambda R_{\langle e \langle et \rangle \rangle} \lambda x. R(y)(x)$

To account for relational uses of nouns that usually have sortal denotations, such as *horse*, I proposed the rule (27a), repeated below in (29). It is essentially the same type-shifting operation proposed in other accounts, such as Barker 2011 and Coppock & Beaver 2015, but with the syntactic effect of producing a transitive noun. In (29), P is the denotation of the noun that undergoes type-shifting. Hence, feeding this rule with the lexical entry of *horse*, (30a), results in (30b).

- (29) REL := $N \Rightarrow N/NP : \lambda P \lambda y \lambda x. P(x) \land \pi(x)(y)$
- (30) a. $N : \lambda x.\mathbf{horse}(x)$ b. $N/NP : \lambda y \lambda x.\mathbf{horse}(x) \wedge \pi(x)(y)$

A derivation built from the assumptions we have discussed so far is given in (31). The possessive morpheme takes a possessor NP as argument and returns a possessor phrase. The latter then requires some relational denotation to which the possessor can be fed. Since *horse* is a sortal noun, the derivation can only proceed if it acquires a relational denotation via type-shifting. Hence, the type-shifter defined in (29) applies, introducing a relational variable to be contextually set.

- (i) a. Jeg liker den ny-e høvding-en i by-en. I like DEF.M.SG new-DEF chief-DEF.PL in town-M.DEF 'I like the new chiefs of the town.'
 - b. Jeg liker den ny-e farge-n på romm-et. I like DEF.M.SG new-DEF color-M.DEF on room-N.DEF 'I like the new color of the room.'

with respect to the possessive relations it allows. Not all possessive constructions are so permissive, however. Adger (2013:68–69) give the examples in (i), from Norwegian, where different prepositions are compatible with different ranges of relations.

| (31) | Derivation of <i>Joana's horse</i> | | | | |
|------|---|---|---|---|-------|
| | Joana | 'S | | | |
| | NP | $(NP/(N/NP)) \setminus NP$ | | horse | |
| | : j | : $\lambda y \lambda R_{\langle e \langle et \rangle \rangle} \lambda x. R(y)(x)$ | | $N: \lambda x.\mathbf{horse}(x)$ | DEI |
| | | NP/(N/NP) | < | N/NP | - KEL |
| | | : $\lambda R \lambda x. R(\mathbf{j})(x)$ | | : $\lambda y \lambda x$.horse $(x) \wedge \pi(y)(x)$ | |
| | $NP: \lambda x.\mathbf{horse}(x) \wedge \pi(\mathbf{j})(x)$ | | | | |

We still have to account for the derivation of possessive NPs containing temporal modifiers. Let us consider the case of *former*, with the lexical entry in (26), repeated as (32).

(32) former :=
$$N/N : \lambda P \lambda t \lambda x. \neg P(t)(x) \land \exists t'[t' \prec t]. P(t')(x)$$

Examples of the application of this modifier to a sortal and to a relational denotation are given in (33). The last step in (33b) involves a rule of forward composition (Ades & Steedman 1982). The definition of forward and backward composition in (34) is adapted from Steedman & Baldridge (2011).⁷

| (33) | a. | Derivation of <i>former mansion</i> , in its sortal use | | | | |
|------|----|---|--|--|--|--|
| | | former | mansion | | | |
| | | $N/N: \lambda P \lambda x.\mathbf{former}(P)(x)$ | $N:\lambda x.$ mansion (x) | | | |
| | | $N: \lambda x. \mathbf{former}(\mathbf{mansion})(x)$ | | | | |
| | b. | Derivation of <i>former wife</i> | znifa | | | |
| | | Jormer | wije | | | |
| | | $N/N : \lambda P \lambda x. \mathbf{former}(P)(x)$ | $N/NP: \lambda y \lambda x. wife_of(y)(x)$ | | | |
| | | $N/NP: \lambda y \lambda x.$ form | $\operatorname{ver}(\operatorname{wife_of}(y))(x) > D$ | | | |

$$\begin{array}{rcl} \text{(34)} & \text{a.} & X/Y:f & Y/Z:g \Rightarrow X/Z:\lambda z.f(g(z)) & (>B) \\ & \text{b.} & Y\backslash Z:g & X\backslash Y:f \Rightarrow X\backslash Z:\lambda z.f(g(z)) & ($$

In (33), I showed the case of nouns being interpreted in what I

⁷I abstract away from the slash-type hierachy used by Steedman & Baldridge to restrict the application of syntactic rules.

assume are their lexical denotations – *mansion* being lexically sortal, and *wife* being lexically relational. When we consider possessed nouns that are lexically sortal, we face a technical problem under the assumption that the type-shifting operation that turns sortal denotations into relational ones can only apply to resolve a type mismatch. As (33a) shows, *former* can directly combine with the sortal denotation provided by the lexical entry of *mansion*, so there is no mismatch. When the resulting phrase, *former mansion*, combines with the possessor, it can then be shifted into a relational denotation, as shown in (35). However, as this example shows, a relation introduced at this point of the derivation is outside the scope of the temporal modifier. This result is a possible reading of *former mansion*, as predicted, but not the most salient one.

(35) $\lambda y \lambda x.$ former(mansion) $(x) \wedge \pi(y)(x)$

This problem disappears once we allow for flexibility in the typing of noun modifiers. More precisely, by assuming that the so-called Geach Rule (van Benthem 1990:117) is available, as expressed in (36), noun modifiers can be mapped to the type N/NP/(N/NP), corresponding to modifiers of relational nouns.

(36) **Geach rule** An expression occurring in any type $\langle a, b \rangle$ may also occur in type $\langle \langle c, a \rangle, \langle c, b \rangle \rangle$ (for any type *c*).

In a left-to-right derivation of *Mary's former mansion*, we have a possessor phrase requiring a relational argument followed by a noun modifier of type N/N. Given the availability of the Geach Rule, this modifier can shift into a modifier of relational nouns, of category N/NP/(N/NP), as shown in (37). The derivation can then proceed by composition of the possessor phrase with the modifier as in (38).

(37) *former* as a modifier of relation nouns $N/NP/(N/NP) : \lambda R_{\langle e, \langle e, t \rangle \rangle} \lambda y \lambda x. \mathbf{former}(R(y))(x)$ (38) Derivation of a possessor phrase with relational *former*: $\frac{Mary's}{NP/(N/NP)} = \frac{1}{NP/(N/NP)} = \frac{1}{\lambda R\lambda x.R(m)(x)} \frac{N/NP/(N/NP) : \lambda R\lambda y\lambda x.former(R(y))(x)}{NP/(N/NP) : \lambda R\lambda x.former(R(m))(x)} > B$

Alternatively, the shifted modifier can first combine with a relational noun denotation, forming a modified relational nominal (39), which can then combine with a possessor.

(39)

 $\frac{\text{mansion}}{N/NP/(N/NP)} = \frac{\text{mansion}}{N:\lambda x.\text{mansion}(x)} \text{REL}$ $\frac{\cdot \lambda R \lambda y \lambda x.\text{former}(R(y))(x)}{N/NP:\lambda y \lambda x.\text{mansion}(x) \wedge \pi(y)(x)} >$

The availability of these alternative derivations, predicting distinct constituency relations, captures the coordination possibilities we find. In (40a), we have a coordination of the non-canonical constituents formed by the possessor phrase and the temporal modifier. In (40b–40c), we have a coordination of modified possessed phrases, the difference between the two cases being whether there is one or two distinct entities related to Maria.

- (40) a. Maria's former and Joana's current mansion.
 - b. Maria's *former mansion* and *current bed and breakfast* is being restored.
 - c. Maria's *former mansion* and *current bed and breakfast* are being restored.

6 Conclusions

In this paper, I defended a two-place approach to the semantics of possessive noun phrases, along the lines of Vikner & Jensen (2002) and Partee & Borschev (2003). In this kind of approach, possessive noun phrases are uniformly headed by a noun denoting a two-place

relation between entities in the domain. This is straightforward in the case of nouns that lexically encode a relation, such as *daughter*, *colony* or *boyfriend*. The role of the possessor phrase in the construction is providing one of the arguments of this relation. When possessive NPs are headed by lexically sortal nouns, such as *platypus*, *table* or *mansion*, we have a type mismatch. This mismatch is resolved by a type-shifting operation that provides a relational denotation for sortal nouns. This operation introduces a free relational variable whose value is contextually set.

I defended this account from recent arguments leveled by Peters & Westerståhl (2013) against a two-place approach to the semantics of possessives. I showed how two-place approaches make use of type-shifting operations that are independently required, even in one-place approaches, to account for non-conventional interpretations of phrases like *John's brother*, where the NP may refer to someone who is not a member of John's family. Moreover, I showed how this approach can account for scope interactions between possessive noun phrases and temporal modifiers that are problematic for oneplace alternatives. Coupled with a flexible syntactic framework, this analysis is also able to derive the correct semantics for cases of nonconstituent coordination in possessive NPs.

Another contribution of this paper lies in its exploration of the difference between syntactic modifiers, like *former*, and morphological ones, like *ex*-. In this particular case, both modifiers have a similar semantics, and both can be interpreted as having scope over the relation between the possessor and the possessee. The main difference between them is that *ex*- is not compatible with the whole range of possible possessive relations. Importantly, I claim that *ex*- cannot modify relations that are not present in the lexical entry of the noun to which it attaches. This result was derived in this paper from the lexical status of the rule introducing *ex*-, under the assumption that free variables cannot be present in lexical entries corresponding to members of open lexical categories. This assumption has the corollary that lexical rules like *ex*- prefixation cannot include free variables in their output, since these have to be valid lexical entries.

In closing this paper, I leave open the urgent task of embedding these results in an explicit theory of the interface between morphology and syntax, and between morphology and semantics, in a categorial grammar framework.

Acknowledgments I would like to thank Bob Frank, Larry Horn, Hadas Kotek, and Zoltán Szabó, for their comments on a preliminary version of this work, as well as the audience at CSSP 2017. Thanks are also due to the editor of this volume and the anonymous reviewer, whose feedback greatly improved this paper. All remaining errors are, of course, my own.

References

- Ades, A., & Steedman, M. (1982). On the order of words. *Linguistics and Philosophy*, 4(4), 517–558.
- Adger, D. (2013). A Syntax of Substance. Cambridge, MA: The MIT Press.
- Aikhenvald, A., & Dixon, R. M. W. (2013). Possession and Ownership: A Cross-Linguistic Typology. Explorations in Linguistic Typology. Oxford University Press.
- Barker, C. (1995). *Possessive Descriptions*. Dissertations in Linguistics. Stanford, CA: CSLI Publications.
- Barker, C. (2011). Possessives and relational nouns. *Semantics: An International Handbook of Natural Language Meaning*, 2, 1109–1130.
- Barker, C., & Dowty, D. (1993). Non-verbal thematic proto-roles. In A. J. Schafer (Ed.) *Proceedings of NELS 23*, (pp. 49–62).
- Coppock, E., & Beaver, D. (2015). Definiteness and determinacy. *Linguistics and Philosophy*, *38*(5), 377–435.
- Dowty, D. (1979). Word Meaning and Montague Grammar: The Semantics of Verbs and Times in Generative Semantics and in Montague's PTQ. Dordrecht: D. Reidel.
- Dowty, D. (2003). The dual analysis of adjuncts/complements in Categorial Grammar. In E. Lang, C. Maienborn, & C. Fabricius-Hansen (Eds.) *Modifying Adjuncts*, (pp. 33–66). Berlin: De Gruyter Mouton, reprint 2013 ed.
- Enç, M. (1986). *Tense without Scope: An Analysis of Nouns as Indexicals*. Bloomington, IN: Indiana University Linguistics Club.
- Lieber, R., & Scalise, S. (2006). The Lexical Integrity Hypothesis in a new

theoretical universe. *Lingue e linguaggio*, 5(1), 7–32.

Löbner, S. (1985). Definites. Journal of Semantics, 4(4), 279–326.

- Partee, B. (2011). Genitives: A case study (Appendix to Theo M.V. Janssen's "Compositionality"). In J. van Benthem, & A. ter Meulen (Eds.) Handbook of Logic and Language, (pp. 541–553). Cambridge, MA: The MIT Press.
- Partee, B., & Borschev, V. (2003). Genitives, relational nouns, and argumentmodifier ambiguity. In E. Lang, C. Maienborn, & C. Fabricius-Hansen (Eds.) *Modifying Adjuncts*, (pp. 67–112). Berlin: De Gruyter Mouton.
- Peters, S., & Westerståhl, D. (2013). The semantics of possessives. *Language*, *89*(4), 713–759.
- Prince, E. F. (1981). Toward a taxonomy of given-new information. In P. Cole (Ed.) *Radical Pragmatics*, (pp. 223–255). New York: Academic Press.
- Pullum, G. K. (1976). The Duke of York gambit. *Journal of Linguistics*, 12(1), 83–102.
- Steedman, M., & Baldridge, J. (2011). Combinatory Categorial Grammar. In R. D. Borsley, & K. Börjars (Eds.) *Non-Transformational Syntax*, (pp. 181–224). Chichester, West Sussex, UK: Wiley-Blackwell.
- Tonhauser, J. (2002). A dynamic semantic account of the temporal interpretation of noun phrases. *Semantics and Linguistic Theory*, *12*(0), *286–305*.
- van Benthem, J. (1990). Categorial Grammar and Type Theory. *Journal of Philosophical Logic*, 19(2), 115–168.
- Vikner, C., & Jensen, P. A. (2002). A semantic analysis of the English genitive: Interaction of lexical and formal semantics. *Studia Linguistica*, *56*(2), 191–226.