



Corpus, experimental and modeling investigations of cross-linguistic differences in pronoun resolution preferences

MIRIAM SCHULZ 

HEATHER BURNETT 

BARBARA HEMFORTH 

**Author affiliations can be found in the back matter of this article*

RESEARCH

]u[ubiquity press

Abstract

We investigate the impact of syntactic alternatives on pronoun resolution in ambiguous constructions in English and French. Previous research detected language-specific preferences in pronoun resolution in utterances of the type “The postman met the streetsweeper before he went home”. These preferences have been attributed to the interaction of information structural and syntactic constraints inducing a subject bias on the one hand, and Gricean reasoning processes taking into account alternative syntactic constructions on the other hand. A corpus study of four English and French corpora shows that an alternative construction which takes a subject antecedent (“The postman met the streetsweeper before going home”) is much less frequent in spoken English than French. A Rational Speech Act (RSA) model with corpus frequencies integrated as language-specific costs on the use of each construction makes empirical predictions for pronoun resolution preferences in French and English for sentences with “avant”/“before” which have been tested before but also for sentences with “après”/“after” which have not been tested so far. New experimental data show a very good fit of the model predictions for pronoun resolution preferences in English as well as for the differences in antecedent choices between French and English. However, experimental data showing differences in antecedent choices between French sentences with “après” and “avant” deviate from model predictions, indicating that more factors need to be taken into account. The combination of Bayesian modeling, corpus analyses and experimental data shows that RSA models can make relevant and falsifiable predictions for cross-linguistic variation in processing.

CORRESPONDING AUTHOR:
Miriam Schulz

Department of Language
Science and Technology (LST),
Universität des Saarlandes,
66123 Saarbrücken, DE
mschulz@coli.uni-saarland.de

KEYWORDS:

alternative constructions;
pronoun resolution; corpus
study; Rational Speech Act
model

TO CITE THIS ARTICLE:

Schulz, Miriam, Heather
Burnett and Barbara Hemforth.
2021. Corpus, experimental
and modeling investigations
of cross-linguistic differences
in pronoun resolution
preferences. *Glossa: a journal
of general linguistics* 6(1):
66. 1–21. DOI: [https://doi.
org/10.5334/gjgl.1142](https://doi.org/10.5334/gjgl.1142)

An important current puzzle in psycholinguistics is how to account for variation in the resolution of ambiguous pronouns across languages. This puzzle can be illustrated by data such as those in (1), from Hemforth et al. (2010).

- (1) a. *English:*
The postman met the streetsweeper before he went home.
- b. *French:*
Le facteur a rencontré le balayeur avant qu' il rentre à la maison.
The postman met the streetsweeper before he went home.
- c. *German:*
Der Briefträger hat den Straßenfeger getroffen bevor er nach Hause ging.
The postman has the streetsweeper met before he home went.

Previous research has shown that, while the sentences are superficially similar, the ambiguous pronoun that is the subject of the subordinate clause in (1) is interpreted differently across languages (Hemforth, Colonna, Pynte & Konieczny 2004; Hemforth, Konieczny, Scheepers, Colonna, Schimke & Pynte 2010; Colonna, Schimke & Hemforth 2012, 2014; and Baumann, Konieczny & Hemforth 2014). The pronoun can refer back either to the subject (the postman) or the object (the streetsweeper) of the matrix clause. A number of experiments in the above-mentioned sources have reliably demonstrated that English and German speakers prefer to resolve the ambiguous pronoun in this kind of sentence as referring to the subject of the main clause, while French speakers tend to interpret the pronoun as referring to the object.

The findings for English and German can be explained by general cross-linguistic preferences in anaphora resolution for the first-mentioned antecedent (see e.g. Gernsbacher & Hargreaves 1988) or for the subject (see e.g. Järvikivi, van Gompel, Hyönä & Bertram 2005). However, these approaches cannot explain why there is an object preference in French. Baumann et al. (2014) give a possible explanation for the French data based on Gricean reasoning and the availability of an alternative grammatical construction to express one of the two possible readings of the sentence: they observe that French grammar also generates the expression in (2), which is identical to (1b), except that the complementizer *que* is substituted by the preposition *de*, the subjunctive form *rentre* by the infinitival form *rentrer*, and the pronoun *il* by PRO which is grammatically bound to the subject of the matrix clause and can therefore only be interpreted as referring to the postman.

- (2) Le facteur a rencontré le balayeur avant de PRO rentrer à la maison.
The postman met the streetsweeper before PRO go.INF home.

French speakers obeying Gricean maxims such as *manner* (*be perspicuous, avoid ambiguity, etc.*) should be more likely to resolve the pronoun in the subordinate clause to the object, in order to avoid ambiguity with (2), which has obligatory resolution to the subject.¹

Although this Gricean analysis accounts for cross-linguistic variation between German and French, it runs into problems when applied to English. English also possesses a non-finite construction corresponding to the French one in (2), which is shown in (3). Nevertheless, English (1a) shows similar subject-oriented anaphora interpretation preferences as German (1c).

- (3) The postman met the streetsweeper before going home.

To account for the difference between English and French, Baumann et al. (2014) propose a frequency-based explanation: in a small-scale corpus study (*Europarl*, Koehn 2005), they observe that the French non-finite construction in (2) is much more frequent than its English counterpart in (3), and hypothesize that English speakers simply do not take the existence of the alternative construction in (3) into account in the process of reference resolution when faced with sentences like (1a). Therefore, according to the frequency-based account, the preferences

¹ One of our anonymous reviewers questioned whether the assumption that the non-finite construction unambiguously refers to the subject can be empirically verified. We report the results of a follow-up study conducted to test this assumption in Appendix D of the supplementary materials.

in favor of the subject or first-mentioned antecedent would remain decisive for determining pronoun reference in English.

Although experimental and corpus data so far support the above line of reasoning, along with similar data collected for Portuguese (Baumann, Konieczny & Hemforth 2014) and Catalan (Mayol & Clark 2010), no formal account of the precise workings of, and interaction between, these different mechanisms appearing to be jointly responsible for the observed cross-linguistic variation in pronoun resolution has been developed so far.² Rational Speech Act (RSA) models (Frank & Goodman 2012) provide a means to formalize core parts of Gricean reasoning using the tools of game theory, information theory and Bayesian inference. Such models allow for explicit predictions to be made about the nature and interaction of the diverse factors that underlie language production and interpretation, and therefore naturally lend themselves to the study of complex psycholinguistic phenomena like pronoun resolution. Conceiving of linguistic communication as social cognition (Goodman & Stuhlmüller 2013), they can be used not only to explain and predict experimental as well as corpus data, but further, to reinforce the cognitive plausibility of existing theoretical approaches.

In this paper, we present new data from cross-linguistic and genre-sensitive corpus studies which not only partially reproduce previous findings, but also draw a more fine-grained portrait of the frequencies and conditions of use of the different types of construction in English and French. This provides a much needed counterpart to the extensive body of experimental data that has been gathered on the use of alternative constructions, and it further allows for the investigation of open issues of previous research, such as the question whether the observed cross-linguistic differences can extend to other connectors, which may further differ in verbal mood (Hemforth et al. 2010), and the hypothesis that the preference for the finite construction in English is limited to spoken registers (Baumann et al. 2014). Based on the results obtained in our corpus study as well as on data from experiments and corpus studies conducted in prior research, we develop and fine-tune RSA models (Frank & Goodman 2012; Goodman & Stuhlmüller 2013) of the pragmatic inferences at play in the resolution of ambiguous non-reflexive pronouns. The models integrate the frequencies of the respective constructions found in the corpus study with general constraints on pronoun resolution in order to predict the language-dependent pronoun resolution preferences. The RSA models make precise predictions for pronoun resolution preferences in English and French not only for sentences with *before/avant*, which have been tested in previous experiments, but also for *after/après*. New experimental data from English and French show a very good fit with model predictions for English as well as for the differences in antecedent choices between French and English.

2 Alternative constructions and Gricean reasoning

A vast body of research in pragmatics and psycholinguistics has shown that the existence of alternative utterances in a language can influence how we interpret its expressions. A classic example of this phenomenon is *scalar implicature*: as discussed by Grice (1975), Horn (1984), and Levinson (2000), among many others, in many situations, listeners will interpret a sentence with the quantifier *some*, such as (4), as equivalent to the sentence with *some but not all* in (5), despite the fact that *some* allows for a *some and possibly all* interpretation in many semantic environments (6).

(4) Mary ate **some** of the cookies.

(5) Mary ate **some but not all** of the cookies.

(6) Did you eat **some** of the cookies? (you must answer *yes* if you ate all of them)

The explanation for the generation of the *not all* inference in (4) within the Gricean tradition is that in interpreting (4) listeners also take into consideration the alternative expression (7), and reason about why the speaker chose to utter (4) instead of (7).

² There is a large amount of research (both on the experimental and on the formal modeling side) into the interpretation of null and overt pronouns in the context of pro-drop (Carminati 2002; Mayol & Clark 2010; de la Fuente et al. 2013, 2016, among many others). In this paper, we focus only on the non-pro-drop languages English, French and German.

(7) Mary ate **all** of the cookies.

In particular, our conversations take place under the understanding that, when we communicate, we will obey certain principles (or *maxims*) such as *Make your contribution as informative as possible (for the current purposes of the exchange)* (**Quantity**); *Make your contribution true* (**Quality**), and *Be perspicuous; so avoid obscurity and ambiguity* (**Manner**). Since (7) is true only in a subset of the situations in which (4) is true (the ones in which Mary ate all the cookies), (7) is more informative. Therefore, given the maxims of *Quantity* and *Quality*, the listener reasons that if (7) were true, the speaker would have uttered it rather than (4). Since they uttered (4), the listener concludes that (7) is not true, i.e. draws the *not all* implicature.

Gricean-style explanations have also been applied to syntactic parsing preferences, and cross-linguistic differences in language processing. For example, although sentences such as those in (8) are, in principle, ambiguous between a parse in which the relative attaches to the direct object (the daughter of the colonel: *high attachment*) and one in which the relative attaches to the complement of daughter (the colonel: *low attachment*), English speakers prefer the low attachment parse, whereas Spanish speakers prefer the high attachment parse (Cuetos & Mitchell 1988, Frazier & Clifton 1996).

- (8) a. The journalist interviewed the daughter of the colonel *who had the accident*.
b. El periodista entrevistó a la hija del coronel *que tuvo el accidente*.

Frazier & Clifton (1996) propose that the observed differences between English and Spanish arise from listeners reasoning about the different sets of alternative constructions available in the two languages. More specifically, unlike Spanish, English has a Saxon genitive construction (9) in which the relative can only attach to the direct object (the colonel's daughter).

(9) The journalist interviewed the colonel's daughter *who had the accident*.

Since (9) can describe only a subset of the situations that (8a) can describe, (9) can be considered more informative than (8). Therefore, English listeners hearing (8a) will be more likely to draw the inference that it was the colonel who had the accident, because if it was the daughter, the speaker would have used the Saxon genitive construction (9). Spanish has no construction comparable to (9); therefore, no alternative-based reasoning takes place and the relative clause most often attaches to the first mentioned constituent, the direct object in this case.

As mentioned in the introduction, Hemforth et al. (2010) and Baumann et al. (2014) provide Gricean-style explanations for cross-linguistic differences between German and French pronoun resolution. French has an alternative non-finite construction which can describe only the situations in which the postman goes home (11),³ and this alternative is more informative than the finite alternative (10b), which can describe both situations: one in which the postman goes home or one in which the streetsweeper goes home. Therefore, through Gricean reasoning, listeners are predicted to prefer to interpret the pronoun *il* in the finite subordinate clause in (10b) as referring to the direct object, which is compatible with the empirical findings. German has no relevant non-finite construction; therefore, no alternative-based reasoning takes place and the pronoun *er* in (10a) is predicted to most often refer to the first mentioned referent (the postman), which is consistent with a general subject or first-mention preference.

- (10) **Finite construction**
a. Der Briefträger hat den Straßenfeger getroffen bevor **er** nach Hause ging.
b. Le facteur a rencontré le balayeur avant qu'**il** rentre à la maison.

³ It has been argued that French as other Romance languages allows for experiencer objects to be antecedents for PRO in sentences like “Le parachutisme effraye Pierre avant même de PRO y avoir été initié.” (“Skydiving scares Peter even before PRO being initiated to it.”; Legendre 1993). A follow-up study (see supplementary materials, Appendix D) shows, however, that subject antecedents are a near categorical choice for French speakers with 96% of subject choices in sentences like (11) with *avant* or *après* (similar to English speakers with 93% subject antecedent choices in non-finite constructions with *before* or *after*), which is within the range of error expected in online experiments (see, for example, Hemforth et al. 2020).

(11) **Non-finite alternative construction**

Le facteur a rencontré le balayeur avant de **PRO** rentrer à la maison.

The accounts of scalar implicature, relative clause attachment and pronoun resolution described above crucially rely on the correct identification of the set of alternatives that are input to Gricean reasoning. However, the precise criteria for defining sets of alternative constructions within a language and cross-linguistically remains an open question in formal pragmatics and psycholinguistics (see Atlas & Levinson 1981; Horn 1989; Chierchia 2004; Katzir 2007; Fox & Katzir 2011, among others). An influential proposal in formal pragmatics is that of Katzir (2007) and Fox & Katzir (2011) who propose that sets of alternatives are generated by three operations on linguistic expressions: deleting constituents, substituting constituents with elements from the lexicon, and replacing constituents with material provided by the context.

Under Fox & Katzir's theory (see also Chierchia 2004), (4) and (7) count as alternatives, since (7) is the result of substituting *all* for *some* in (4); (8a) and (9) count as alternatives, since (9) can be derived from (8a) by substituting the case marker 's for the preposition *of* (and satisfying the syntactic requirements of 's); and (10b) and (11) count as alternatives, since (11) can be derived from (10b) by substituting the complementizer *que* for the preposition *de*, the pronoun *il* by **PRO**, and the subjunctive form *rentre* by the infinitival form *rentrer*.

Although alternative-based reasoning accurately predicts variation in pronoun resolution between French and German, as mentioned in the introduction, the simple existence of an alternative (as defined by Fox & Katzir) in the language does not always appear to trigger Gricean reasoning. English (12a) and (12b) are alternatives in the Fox & Katzir sense, since (12b) can be derived from (12a) by substituting *going* for *went* and **PRO** for *he*; however, English speakers still show a subject preference for the interpretation of *he* in (12a).

- (12) a. The postman met the streetsweeper before **he** went home. **Finite**
b. The postman met the streetsweeper before **PRO** going home. **Non-finite**

Nevertheless, Hemforth et al. (2010) and Baumann et al. (2014) observe that the French and English finite and non-finite constructions are not identical in every way. Baumann et al. (2011, 2014) show in a corpus study that the non-finite construction is over 1.5 times more frequent in French than the finite construction, while in English, it is the finite construction which is more than four times as frequent as its non-finite counterpart. They therefore hypothesize that **frequency** also plays a role in defining the set of alternatives triggering Gricean reasoning, and that the English non-finite construction (12b) is simply not frequent enough to be accessible in language processing and thus to count as an alternative to (12a).

The notion that frequency plays a role in the definition of the set of alternatives, and that it can explain cross-linguistic variation between French and English, is appealing; however, many of its aspects and consequences are currently under-developed. For example, it remains an open question where such vast differences in corpus frequencies of parallel constructions among languages come from. In a side note, Baumann et al. (2011: p. 3297, 2014: p. 208) propose the hypothesis that the non-finite construction in English may be preferably used in written registers as compared to spoken language, which could in turn explain the diminished impact of the alternative construction on ambiguous pronoun resolution. So far, genre-, modality- or register-specific variation of the frequency of alternative constructions has not been empirically explored. Should the alternative construction prove to be mostly confined to written registers in English, this could explain why English speakers, as opposed to French speakers, do not take it into account in pronoun resolution. In order to test this hypothesis, as well as the general impact of modality and genre, we conducted corpus analyses across three genres (spoken, newspaper, literature) in English and French to provide a more detailed picture of the distribution of finite and non-finite constructions for *before/avant* and *after/après*. The individual frequencies of each construction in the spoken corpora are then integrated as costs into the Rational Speech Act models developed for each language and connector in Section 4, whose precise quantitative predictions are in turn evaluated against new empirical data in Section 5.

3 Corpus study

The corpus analysis presented in this section further extends the domain of alternatives under investigation to complementizer and prepositional uses of *after*, which provide a further pair of alternative constructions consisting of an ambiguous finite construction and a non-finite counterpart in which the zero anaphor is bound to the subject of the matrix clause (13)–(14).

- (13) **English**
- a. The postman met the streetsweeper after he went home.
 - b. The postman met the streetsweeper after going home.
- (14) **French**
- a. Le facteur a rencontré le balayeur après qu'il est rentré à la maison.
 - b. Le facteur a rencontré le balayeur après être rentré à la maison.

An overview of the different types of the pairs of alternative constructions in English and French – anaphoric embedded clauses with *before/avant* and *after/après* – that constitute the object of inquiry of the following corpus study are presented in [Table 1](#), along with their relevant formal properties. Note that, unlike in English, the prepositional uses of *avant que* and *après que* differ in the mood of the embedded clause.

Construction type	English	French
finite construction	<i>before</i> + finite phrase	<i>avant que</i> + finite phrase in subjunctive mood
alternative non-finite construction	<i>before</i> + gerund phrase	<i>avant de</i> + infinitive
finite construction	<i>after</i> + finite phrase	<i>après que</i> + finite phrase in indicative mood
alternative non-finite construction	<i>after</i> + gerund phrase	<i>après</i> + infinitive of <i>avoir/être</i> + past participle

Table 1 Overview of the different types of construction investigated in the corpus study.

Hemforth et al. (2010) advanced the hypothesis that the variation between subjunctive and indicative mood following different French conjunctions provides an obstacle for French speakers using the finite construction with respect to English speakers. French speakers often use the indicative mood after *avant que* (Kastronic 2016), despite the fact that normative grammarians prescribe the subjunctive (Poplack et al. 2013). However, the conjunction *après que* has seen the inverse effect: while normative grammarians continue to insist that it be followed by indicative mood, it is frequently employed with the subjunctive mood, especially in spoken French (Canut & Ledegen 1998; Kastronic 2016). It is possible that uncertainty surrounding the use of the conjunctions takes its toll on the ease with which the finite construction is produced. This could then help explain the French preference for the non-finite construction for which there is no need to decide which mode to employ, and conjugation is dispensed with altogether. Furthermore, the production of the subjunctive may be considered more cognitively taxing than that of the indicative. It is therefore also interesting to investigate the finite and non-finite uses of *après* alongside *avant*. Our corpus study thus investigates the uses of conjunctive vs. prepositional *after* or *après* and compares the respective uses of both connectors with each other. Results from the corpus analysis are used for a Rational Speech Act model that predicts preferences for antecedent choices in experimental data. Those predictions are then tested in questionnaire studies in English and French.

3.1 The corpora

In prior corpus studies, Baumann et al. (2011, 2014) made use of *Europarl* (Koehn 2005), a parallel corpus which consists of a collection of the proceedings of the European parliament translated into each of the 11 official languages of the European Union. However, the fact that *Europarl* consists of translations means that the bias of the translator could constitute a factor influencing the results, and furthermore, it is restricted to a single genre and thus does not offer the diversity of genres required for the present purposes. Therefore, for the English study,

we choose to use the *Corpus of Contemporary American English* (COCA, Davies 2008). COCA currently consists of over 570 million part-of-speech tagged words, with 20 million words added each year evenly divided across 5 different genres since 1990. The fact that its data are spread evenly across different genres and modalities and that the data within each genre are further balanced across numerous sub-genres makes it a natural candidate for the present purposes. The relevant sections of COCA include *Fiction* (over 111 million words), *Newspaper* (over 112 million words) and *Spoken* (over 116 million words). The sources within each section are diverse, with texts in the *Fiction* section ranging from plays to magazines, the *Newspaper* section including articles from 10 newspapers from different domains, and the *Spoken* section consisting of unscripted conversation from over 150 different radio and TV programs.

No single resource exists for French that would be comparable to COCA in terms of size or variety of genres. We therefore studied multiple French corpora, which were selected based on their size and their correspondence to the different sections of COCA. *Frantext* (<https://www.frantext.fr/>), a corpus containing 251 million words belonging to diverse literary genres, was chosen as a rough counterpart to COCA's *Fiction* section. From 2018 onwards, the renewed online query interface of *Frantext* allows to formulate searches using regular expressions and a generalized Corpus Query Language (CQL). However, as the sources of *Frantext* date back many centuries, a filter was added to restrict the date range of the sources to those from 1990 and later to be on a par with COCA in order to control for effects of diachronic change. This resulted in a sub-corpus containing over 22 million words.

The *Est Républicain* corpus provides a large corpus of journalistic French and can therefore constitute a counterpart to COCA's *Newspaper* section. It is freely available for download and could thus be interrogated by advanced tools for corpus linguistics such as *AntConc* (Anthony 2018) using regular expressions. The corpus contains roughly 149 million words collected from the *Est Républicain*, a local newspaper based in the eastern part of France, over the years 1999, 2002 and 2003 (Seddah et al. 2012; ATILF & CELLE 2020).

Finally, a lemmatized version of the two *ESLO* corpora (*Enquête sociolinguistique à Orléans*, <http://eslo.huma-num.fr/>) provides a rich resource of transcriptions of spoken French containing over 4 million words collected from sources ranging from phone conversations to public debates from 1969 to 1974 (*ESLO1 corpus*) and 2008 to 2012 (*ESLO2 corpus*). With the diversity of the sources the *ESLO* corpora contain, they constitute a fairly balanced corpus. Their focus on daily interactions and limitation to speakers from mostly the French city of Orléans nevertheless slightly set them apart from COCA's more media-focused and demographically balanced *Spoken* section.

3.2 Procedure

We restricted our search to non-cataphoric conjunctive and prepositional uses of *before* (*avant*) and *after* (*après*). A full list of the CQL expressions and corpus hits obtained is detailed in Appendix A. English non-finite alternative constructions were required to contain a gerund, French non-finite alternatives an infinitive. Only cases where the constructions did not appear in sentence-initial position were taken into account,⁴ since cataphoric pronouns are subject to different constraints than anaphoric pronouns (Fedele & Kaiser 2014).

3.3 Results

Tables 2 and **3** show the raw counts of sentence-final occurrences for each of the three sections of COCA together with the proportion of the alternative (non-finite) construction with respect to the finite construction. **Table 2** summarizes the results for *before*, while **Table 3** displays those for *after*.

First of all, these results show that the alternative construction is generally much less frequent in both spoken English and literary English for both connectors, while the frequency ratios are reversed in journalistic English where the alternative construction is used more frequently than the finite construction, especially with prepositional *after*.

⁴ Except for the *ESLO* corpora which contain no marking of sentence boundaries; here, we manually checked that no match for any of our queries occurred at the beginning of a new speaker turn.

	Spoken	Fiction	Newspaper
Finite construction	25,087	39,309	14,897
Alternative (non-finite) construction	3,065	13,047	16,613
Proportion non-finite	0.11	0.25	0.53

	Spoken	Fiction	Newspaper
Finite construction	12,930	15,058	12,244
Alternative (non-finite) construction	8,523	7,846	27,021
Proportion non-finite	0.40	0.34	0.69

As expected, the French corpora draw a very different picture. *Tables 4* and *5* show the number of corpus hits for sentence-final constructions with *avant* and *après*. The number of raw corpus hits shown in the tables differs drastically depending on corpus size because the French corpora are not as balanced with respect to size as the sections of *COCA*. Nevertheless, the proportion of the alternative construction compared to the finite construction shows the relative distribution of the two constructions across the respective corpus.

	ESLO	Frantext	Est Républicain
Finite construction	88	848	692
Alternative (non-finite) construction	241	5,021	22,382
Proportion non-finite	0.73	0.86	0.97

	ESLO	Frantext	Est Républicain
Finite construction	22	274	178
Alternative (non-finite) construction	73	2,358	12,040
Proportion non-finite	0.77	0.90	0.99

Across the French corpora, the results also show great differences in the use of the alternative construction with respect to the finite construction for both *before* and *after*. Although it is used more frequently than the finite construction across all three genres, the alternative construction is used even more frequently in journalistic French with respect to spoken or literary French. In fact, French shows the same tendency here as English towards highly increased use of the alternative construction in journalistic language. With the exception of the data for *after* in English (see *Table 3*), the alternative construction is furthermore used more often in literary than in spoken language in both English and French. *Figures 1* and *2* show the corpus results as a proportion of non-finite constructions for French and English respectively.⁵

For inferential statistics, we coded finite constructions as zero and non-finite constructions as 1 to serve as the dependent variable. The predictors (connector, language) were mean centered coded. The predictor *genre* was mean centered coded with *spoken* as the reference category to which *literary* and *news* are compared.⁶ The predictors as well as all interactions were submitted to a logistic regression model (see *Table 6*, as well as Appendix C for the logistic regression formula).

⁵ Error bars are calculated based on normal approximation of the binomial distribution which is adequate for our sample sizes.

⁶ Predictors are coded such that positive estimates mean for connector: more non-finite constructions for *après/after* than *avant/before*; for language: more non-finite constructions for *French* than for *English*; for literary-spoken: more non-finite constructions for *literary* than for *spoken*; for news-spoken: more non-finite constructions for *news* than for *spoken*.

Table 2 Corpus counts of sentence-final *before* in English.

Table 3 Corpus counts of sentence-final *after* in English.

Table 4 Corpus counts of sentence-final *avant* in French.

Table 5 Corpus counts of sentence-final *après* in French.

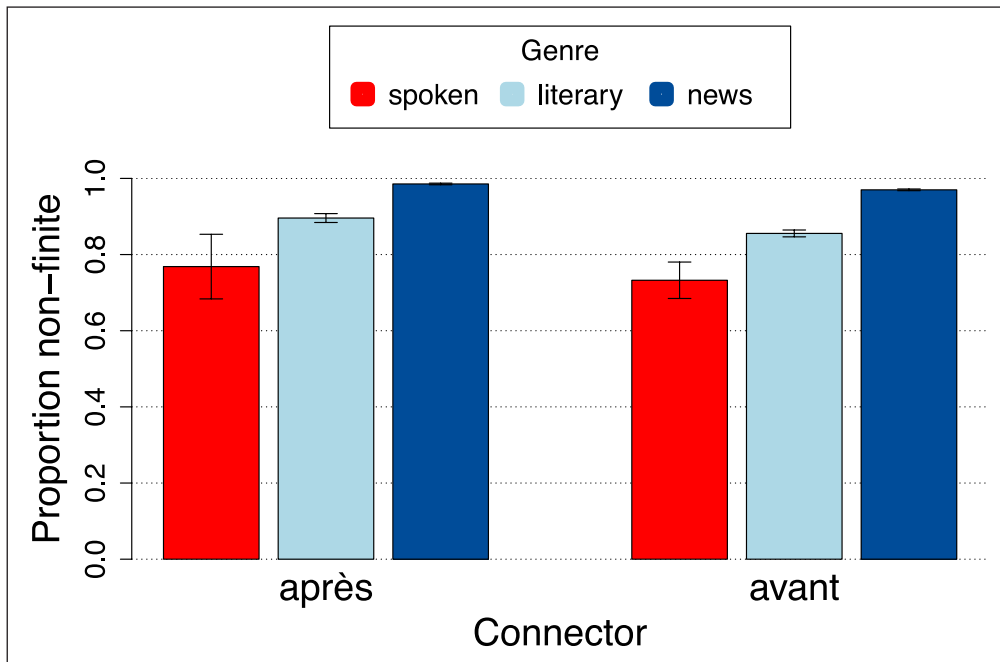


Figure 1 Proportion of non-finite constructions for *après* (after) and *avant* (before) in the French corpora.

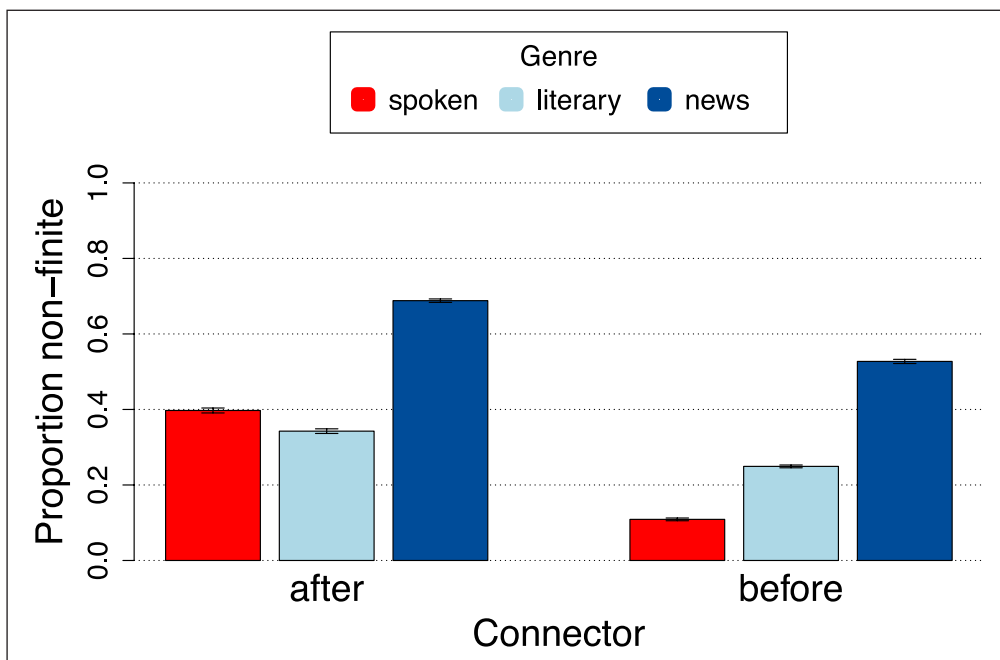


Figure 2 Proportion of non-finite constructions for *after* and *before* in the English corpora.

As shown in [Table 6](#), all predictors have significant main effects with significantly fewer non-finite constructions for *before/avant* than for *after/après*, significantly more non-finite constructions in news corpora as well as in literary corpora compared to spoken corpora, and finally significantly more non-finite constructions in French than in English.

The central outcome for this paper is, however, the significant three-way interaction between connector, language and literary and news compared to the spoken genre, which we later investigate in a little more detail. Given that the written corpora (news and literary) are subject to editing to varying extents and may thus reflect stronger adherence to linguistic norms, we focus on the data from spoken corpora to derive model predictions.

A more general pattern in the corpora that we analyzed concerns the distribution of the connectors *after/après* and *before/avant* across genres and languages. Logistic regressions predicting connector type with *after/après* coded as 0 and *before/avant* coded as 1 as dependent variable and genre and language as mean centered predictors (with *spoken* as reference category for genre) show that *before/avant* constructions are generally preferred across languages (Estimate = 0.353579, Std. Error = 0.006171, $z = 57.298$, $p < .0001$) with significantly

	Estimate	Std. Error	Z-value	p-value
intercept	0.113588	0.007477	15.192	<.0001
connector	0.752998	0.015906	47.342	<.0001
literary-spoken	0.557299	0.026717	20.860	<.0001
news-spoken	1.963992	0.026866	73.103	<.0001
language	3.014905	0.033586	89.768	<.0001
connector:literary	-0.973401	0.057382	-16.963	<.0001
connector:news	-0.717452	0.057606	-12.455	<.0001
connector:language	-0.314011	0.073668	-4.263	<.0001
literary:language	0.353739	0.129486	2.732	<.01
news:language	0.894099	0.130672	6.842	<.0001
connector:literary:language	1.416511	0.284561	4.978	<.0001
connector:news:language	1.548790	0.287504	5.387	<.0001

Table 6 Logistic regression results for the corpus study.

more *before/avant* constructions in French than in English (Estimate = 0.572198, Std. Error = 0.026536, $z = 21.563$, $p < .0001$). However, two-way interactions between language and genre show that both languages show a similar preference for *before/avant* constructions in the *literary* genre while the higher proportion of *before/avant* in French holds for the *news* and for the *spoken* genre (literary-spoken*language: Estimate = -0.995242, Std. Error = 0.119420, $z = -8.334$, $p < .0001$; news-spoken*language: non significant, $p > .3$). *After* and *before* are more evenly distributed in English than in French for these two genres (see [Figure 3](#)). We discuss the possible implications of this finding in Section 5.

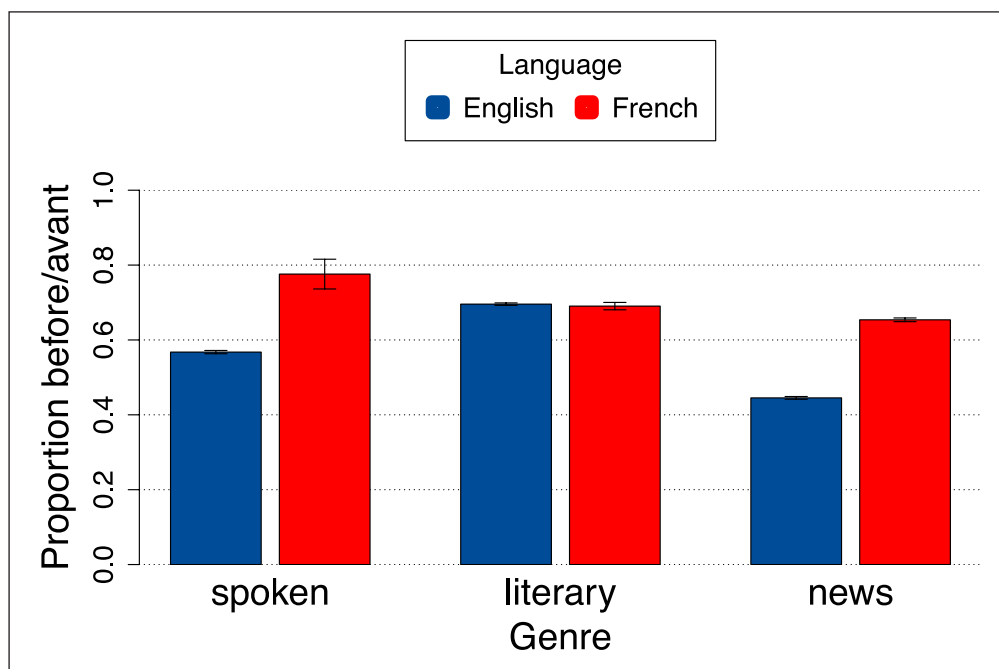


Figure 3 Proportion of *before/avant* constructions (out of *before/avant* + *after/après*).

3.4 Discussion

While many highly interesting conclusions can be drawn from the corpus data presented in this section,⁷ we simply highlight those which are of central importance for testing our hypotheses using formal modeling in the next section:

⁷ For more detailed discussion and explanation of genre-related differences in frequencies of *before/after* see Schulz (2018).

1. The results of the present study confirm the general preference for the finite construction in English and for the alternative construction in French obtained by Baumann et al. (2011; 2014) although the preferences depend on the genre with generally more non-finite constructions in the news corpora.
2. While the non-finite construction is equally prevalent for *après* and *avant* in French, it is significantly less frequent for *before* than for *after* in English.
3. *Before/avant* constructions are more frequent in literature across languages as well as in spoken and in news corpora in French, but not in English.

We now incorporate these findings into *Rational Speech Act* models in order to identify the predictions that a Gricean Reasoning + Frequency-based theory makes for pronoun resolution with *avant*, *après*, *before* and *after*.

4 A Rational Speech Act model of cross-linguistic pronoun resolution

The results of the corpus study in Section 3 further reinforce the question if overall frequencies in particular in spoken language could condition whether an utterance can be perceived as an alternative to another utterance and as such, impact the interpretation of an ambiguous pronoun in the latter. However, it is not clear how exactly to integrate this observation into the Gricean analysis for pronoun resolution sketched in Section 2. Furthermore, as has been frequently observed since Kroch (1972), Horn (1972) and Gazdar (1979), Grice's maxims, as stated, need to be refined in order to make clear predictions for the wide range of pragmatic phenomena that we find in natural languages. Grice (1975) himself viewed linguistic communication as an instance of more general rational behaviour, saying (p. 47),

As one of my avowed aims is to see talking as a special case or variety of purposive, indeed rational, behaviour, it may be worth noting that the specific expectations or presumptions connected with at least some of the foregoing maxims have their analogues in the sphere of transactions that are not talk exchanges.

Building on this idea, the field of *game-theoretic pragmatics* uses the tools of *game theory*, a formal theory of interaction (see Osborne & Rubinstein 1994 for an introduction), to model the interactive decision making processes involved in pragmatic inferences (see Benz et al. 2005; Franke 2009, among others). Game-theoretic frameworks, such as the *Rational Speech Act* model (RSA, Frank & Goodman 2012), provide a precise mathematical framework formalizing key aspects of Gricean pragmatics, which allows for the integration of the corpus frequencies observed in Section 3. The RSA architecture is based on a *signaling game* (Lewis 1969). In a signaling game, there are two players: the speaker (S) and the listener (L). S observes a fact about the world, called their *type*, and their goal is to communicate this fact to L. S has a set of *messages*, linguistic forms paired with semantic meanings, which they can use to help communicate their type to L. S's action is to pick a message to send to L. L's action is to hear S's message and to assign it some interpretation and, in doing so, update their beliefs about the world. If S's type and L's message interpretation coincide, then both players win. Otherwise, they both lose. What distinguishes the RSA (and similar frameworks, see Franke 2009; Qing & Franke 2015) from other approaches in game-theoretic pragmatics is 1) its use of an *iterated best response* solution concept which models the 'back and forth' reasoning of speakers and listeners about their interlocutors' behaviour, 2) its formalization of Gricean *Quality* and *Quantity* as *conditionalization* and *negative surprisal* respectively, and 3) its characterization of listener interpretation as *Bayesian inference*. We illustrate how RSA models work by building one for pronoun resolution in French. We then build corresponding models for German and English and discuss the predictions that they make for cross-linguistic variation.

4.1 RSA models of pronoun resolution in French

An RSA model consists of two agents, S and L, a set of possible worlds W , a set of messages M , an interpretation function $[[\cdot]]$, a cost function C , and a function representing L's prior beliefs Pr . In this section, we describe these different components for our proposed models for pronoun resolution after *avant* and *après* in French. All the calculations presented in this paper were

4.1.1 Modeling pronoun resolution in subordinate clauses introduced by *avant*

W represents the range of different possible worlds that S could be communicating to L . For modeling the interpretation of the pronoun *il* in *Le facteur₁ a rencontré le balayeur₂ avant qu'il rentre à la maison*, there are two relevant states of affairs: w_1 , where the mailman (NP_1) went home, and w_2 , where the streetsweeper (NP_2) went home; so $W = \{w_1, w_2\}$. As discussed above, French possesses a non-finite alternative construction ... *avant de PRO rentrer à la maison* which can be used to describe only w_1 . Therefore, $M = \{\textit{avant qu'il rentre à la maison}, \textit{avant de PRO rentrer à la maison}\}$, and the interpretation function $[[.]]$ is defined as in (15).

- (15) a. $[[\textit{avant qu'il rentre à la maison}]] = \{w_1, w_2\}$
 b. $[[\textit{avant de rentrer à la maison}]] = \{w_1\}$

Since we test our models using experimental data in Section 5, the corpus frequencies that are most relevant for the rest of the paper are the ones from spoken corpora. In order to integrate the frequencies into the model, we encode them as part of the *cost* function, C , by taking the natural logarithm of the proportions in the corpus (for French, see [Table 4](#)). Since in the ESLO corpus, the proportion of use of *avant de rentrer à la maison* is 0.73, meaning that the proportion of *avant qu'il rentre à la maison* is 0.27, $C(\textit{avant de rentrer à la maison}) = -0.315$ and $C(\textit{avant qu'il rentre à la maison}) = -1.309$, as shown in [Table 7](#).

	Proportion in spoken corpus P	Message cost = $\log(P)$
Finite construction	0.27	-1.309
Alternative non-finite construction	0.73	-0.315

Table 7 Transformation from corpus proportions to message costs for French *avant*.

The final component of the RSA model is Pr : a specification of the listener's hypotheses concerning which state of affairs is most likely at the point when they hear the main clause *Le facteur a salué le balayeur*, but prior to their hearing the subordinate clause *avant que/de...* We assume that general sentence-level constraints on pronoun resolution such as subject preference, syntactic parallelism, and the impact of linear order shapes the prior. In particular, many continuation studies of pronoun resolution have shown that, in continuations with sentences with two NPs like the ones we are modeling, when speakers use a pronoun, it refers to the subject NP around 80% of the time (80% in Arnold 2001; 78.5% in Kehler & Rhode 2019, for example). We therefore assume that $Pr(w_1) = 0.8$ and $Pr(w_2) = 0.2$.

As the first step in their interpretation process, the listener in an RSA model reasons that their interlocutor is obeying Grice's maxim of *Quality* when they choose their utterance. The RSA formalization of *Quality* is done through proposing that, when they hear the message, the listener *conditions* their prior beliefs under the meaning of the message (16). In other words, as shown in (16), they exclude from consideration all the worlds that are not in the meaning of the message and then normalize the measure.

- (16) For all $w \in W$ and $m \in M$,

$$Pr(w|[[m]]) = \frac{Pr(\{w\} \cap [[m]])}{Pr([[m]])}$$

Since the meaning of *avant qu'il rentre à la maison* includes both worlds, conditionalization just returns the prior probabilities, as shown in [Table 8](#). However, since the meaning of *avant de rentrer à la maison* contains only w_1 , after hearing the alternative non-finite construction, the listener discards w_2 and puts all the probability mass on w_1 .

Grice's maxim of *Quantity* is formalized through the speaker utility function: Frank & Goodman (2012) propose that the utility for a speaker S to use a message m is partially determined by m 's *informativity*, which following Shannon (1948), they take to be the *negative surprisal* (natural

⁸ Our code is available at: https://github.com/miriamschulz/RSA_pronoun_resolution.

	w_1	w_2
Finite construction	0.8	0.2
Alternative non-finite construction	1	0

log) of the prior conditioned under the meaning of the message ($Pr(w|[[m]])$) (17). Furthermore, the costs of the message, which are negative in our model, are added to its informativity.

(17) Speaker utility to use a message m to communicate a world w

$$U_s(m, w) = \log(Pr(w|[[m]]) + C(m)$$

	w_1	w_2
Finite construction	-1.532	-2.919
Alternative non-finite construction	-0.315	$-\infty$

As shown in [Table 9](#), $U_s(\text{finite}, w_1) = \log(0.8) + \log(0.27) = -1.532$, $U_s(\text{non-finite}, w_1) = \log(1) + \log(0.73) = -0.315$, and so on.⁹ Observe that the non-finite construction is predicted to be more useful than its finite alternative to communicate w_1 ; however, the reverse is the case to communicate w_2 .

In RSA models, it is assumed that speakers are **approximately rational** (Anderson 1991) when they choose which message to say, that is, they are rational in the sense that their action selection is guided by utility; however, they may not always choose the most useful action. Mental computations can be subject to time and resource limitations, and there may be a certain amount of inherent variability in the system. Therefore, the speaker's probability of using a message m to communicate a world w ($P_s(m|w)$) is given by the *Softmax* choice rule (Luce 1959; Sutton & Barto 1998), shown in (18). The Softmax choice rule makes reference to a parameter, α , which describes the amount of non-determinism in the system. Setting α to ∞ recovers deterministic choice; however, anything less than ∞ predicts variation.

(18) Speaker probability (Softmax choice)

$$P_s(m|w) \propto \exp(\alpha \times U_s(m, w))$$

We use Bayesian parameter estimation with Markov-chain Monte Carlo (MCMC) methods implemented in *WebPPL* to estimate the optimal value for α for our data (based on Appendix IV in Scontras, Tessler & Franke 2017). This method yields values for α between 0.9 and 1 as the best fits for our results, peaking at 0.93 (see [Figure 4](#)). We therefore adopt an α of 0.93.

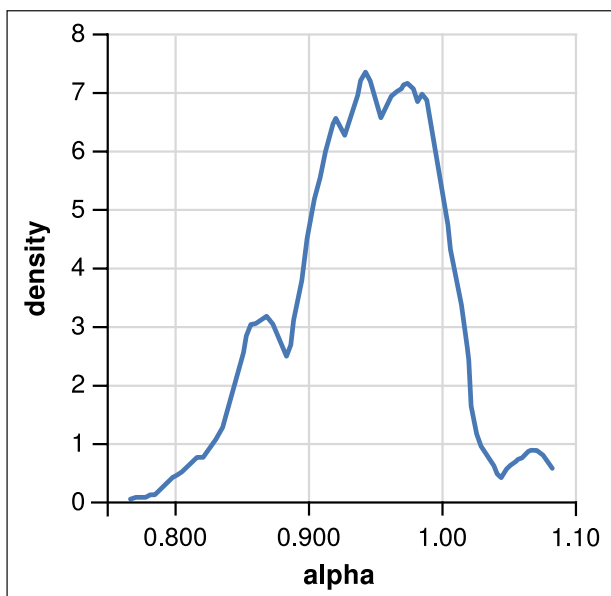


Table 8 Prior conditioned under the meaning of the message ($Pr(w|[[m]])$).

Table 9 Speaker utility to use a message m to communicate a world w ($U_s(m,w)$).

⁹ Note that we numerically define $\log(0)$ as a large negative real number here, given that $x \in [0, 1]$ and $\lim_{x \rightarrow 0^+} \log(x) = -\infty$.

Figure 4 Bayesian parameter estimation for α with MCMC sampling, samples = 10 000, burn in = 2000.

The listener’s final interpretation is given by Bayesian inference: L’s probability of concluding that world w is true, given that they heard the message m , $P_L(w|m)$, is proportional to their prior beliefs in w times the probability that the speaker would use m to signal that world w is true.

$$(19) \quad P_L(w|m) \propto Pr(w) \times P_s(m|w)$$

We can now observe the predictions of the model for pronoun resolution with French *avant* in the finite construction, shown in (20).

- (20) Model predictions for French *avant* (with $\alpha = 0.93$)
- a. $P_L(w_1|finite) = 0.49$ Subject interpretation
 - b. $P_L(w_2|finite) = 0.51$ Object interpretation

We test these predictions in greater detail with new experimental data in Section 5.

4.1.2 Modeling pronoun resolution in subordinate clauses introduced by *après*

The model we build for *après* is structurally identical to the one for *avant* described above. The main difference is the cost associated with the finite and non-finite forms. Since the proportion of non-finite constructions with *après* is 0.77 (recall [Table 5](#)), the cost for the finite construction with this connector is $\log(0.77) = -0.261$. Correspondingly, the cost for the finite construction is $\log(0.23) = -1.470$. The differences in the frequency-induced costs associated with messages generate slightly different predictions for pronoun resolution in finite constructions with *après*: we predict a slightly higher probability of reference to the second NP. More technically, in the *après* model, $P_L(w_1|finite) = 0.46$ and $P_L(w_2|finite) = 0.54$.

A detailed comparison of the descriptions and predictions of the models for *avant* and *après* is shown in [Table 10](#).

	French <i>avant</i>	French <i>après</i>
Messages	Finite (<i>avant que...</i>) Non-finite (<i>avant de...</i>)	Finite (<i>après que...</i>) Non-finite (<i>après AVOIR/ÊTRE</i>)
Costs	C(finite) = -1.309 C(non-finite) = -0.315	C(finite) = -1.470 C(non-finite) = -0.261
Prior beliefs	$Pr(w_1) = 0.8$ $Pr(w_2) = 0.2$	$Pr(w_1) = 0.8$ $Pr(w_2) = 0.2$
α	0.93	0.93
Predictions for pronoun resolution in the finite construction	$P_L(w_1 finite) = \mathbf{0.49}$ (subject) $P_L(w_2 finite) = \mathbf{0.51}$ (object)	$P_L(w_1 finite) = \mathbf{0.46}$ (subject) $P_L(w_2 finite) = \mathbf{0.54}$ (object)

Table 10 Comparison of models for French *avant* and *après*.

4.2 RSA models of pronoun resolution in English and German

The models for English *before* and *after* are parallel to those for *avant* and *après*, only the costs for the (non)finite alternatives differ. As discussed in Section 3, the proportions of use of the finite and non-finite constructions in English are very different from those observed for French. In particular, the proportion of use of the English finite construction with *before* is 0.89, so the cost of this message is $\log(0.89) = -0.117$. The proportion of the English finite construction with *after* is 0.60, so its cost is $\log(0.60) = -0.511$. The consequence of these differences is that the models for English predict a subject pronoun resolution preference for both *before* and *after* ($P_L(w_1|finite) = 0.77$ and 0.68 , respectively). The English models are summarized in [Table 11](#).

Since there is no non-finite alternative for German *bevor*, the RSA model for German is much simpler: the single alternative has no costs, and the result is that the listener interpretation function P_L just reproduces the subject preference in the prior. The German model is shown in [Table 12](#).

	English before	English after
Messages	Finite (<i>before he went...</i>) Non-finite (<i>before going...</i>)	Finite (<i>after he went...</i>) Non-finite (<i>after going...</i>)
Costs	C(finite) = -0.117 C(non-finite) = -2.207	C(finite) = -0.511 C(non-finite) = -0.916
Prior beliefs	$Pr(w_1) = 0.8$ $Pr(w_2) = 0.2$	$Pr(w_1) = 0.8$ $Pr(w_2) = 0.2$
α	0.93	0.93
Predictions for pronoun resolution in the finite construction	$P_L(w_1 finite) = 0.77$ (subject) $P_L(w_2 finite) = 0.23$ (object)	$P_L(w_1 finite) = 0.68$ (subject) $P_L(w_2 finite) = 0.32$ (object)

Table 11 Comparison of models for English *before* and *after*.

	German bevor	German nachdem
Messages	Finite (<i>bevor er...</i>)	Finite (<i>nachdem er...</i>)
Costs	C(finite) = 0	C(finite) = 0
Prior beliefs	$Pr(w_1) = 0.8$ $Pr(w_2) = 0.2$	$Pr(w_1) = 0.8$ $Pr(w_2) = 0.2$
α	0.93	0.93
Predictions for pronoun resolution in the finite construction	$P_L(w_1 finite) = 0.8$ $P_L(w_2 finite) = 0.2$	$P_L(w_1 finite) = 0.8$ $P_L(w_2 finite) = 0.2$

Table 12 Model for German *bevor* und *nachdem*.

In summary, the predictions of the RSA models for French and English *avant/before* and *après/after* are shown in [Figure 5](#). In the next section, we test these predictions with new experimental data.

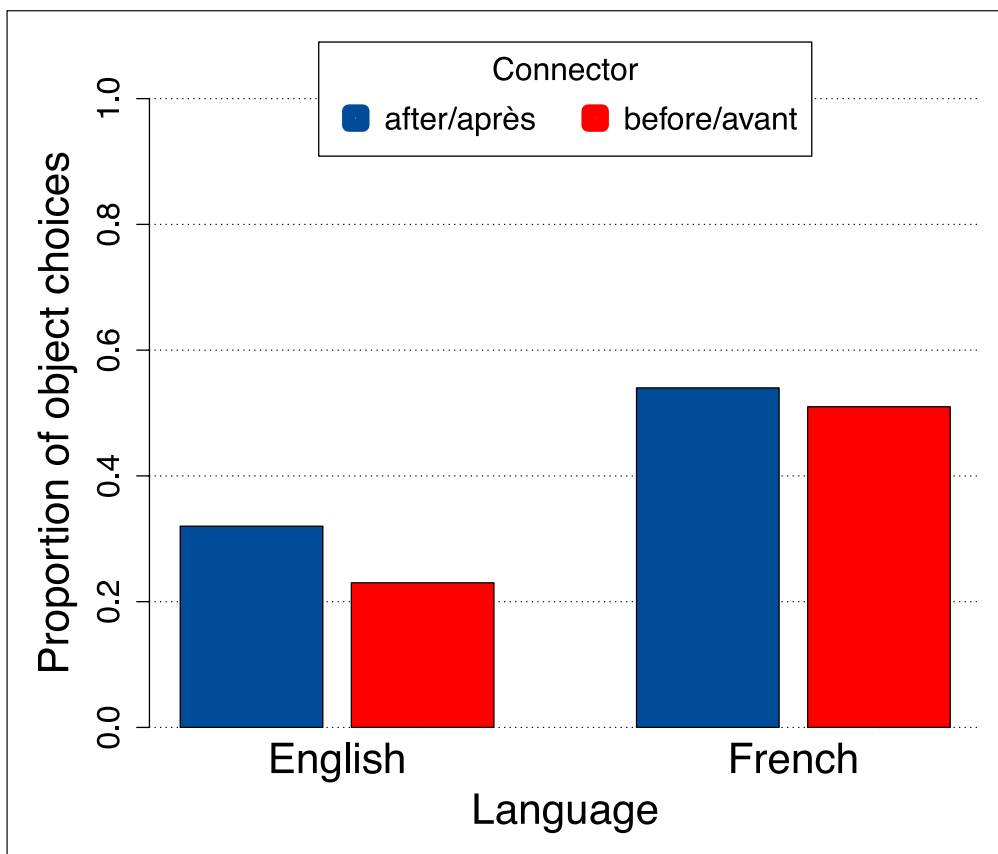


Figure 5 Model predictions for proportions of object choices for the connectors *after/après* and *before/avant* for English and French.

We tested the model predictions in a crosslinguistic *Cloze* experiment similar to Hemforth et al. (2010). Unlike this earlier research, we did not include a cross-sentence condition like (21) which has a strong subject preference across languages. Parallel to the corpus data, we included two connector conditions *before/avant* and *after/après* (see (22) for English and (23) for French). Materials were adapted from Schimke et al. (2018).¹⁰ Subordinate clauses were constructed such that they were equally plausible with a subject and an object antecedent following the intuitions of all authors. To control for possible remaining biases, we added conditions (22c,d; 23c,d) with inverted roles.

- (21) The postman met the streetsweeper. Then he went home.
- (22)
- The policeman called the postman before he tied his shoelaces.
 - The policeman called the postman after he tied his shoelaces.
 - The postman called the policeman before he tied his shoelaces.
 - The postman called the policeman after he tied his shoelaces.
- (23)
- Le policier a appelé le facteur avant qu'il noue ses lacets.
 - Le policier a appelé le facteur après qu'il a noué ses lacets.
 - Le facteur a appelé le policier avant qu'il noue ses lacets.
 - Le facteur a appelé le policier après qu'il a noué ses lacets.

Participants: 83 French and 37 English speakers participated in the experiment (54 female for French, mean age 37 years, min 18 years, max 79 years; 14 female for English, mean age 36 years, min 20 years, max 64 years). French participants were recruited via the RISC website (<https://www.risc.cnrs.fr>) and via social networks, English speakers were recruited via Amazon Mechanical Turk and paid 2\$ for their participation. All speakers indicated either French or English as their native language.

Materials: 16 items following the template in (22) and (23) were created and distributed across four lists following a Latin Square design. The experimental items were mixed with 40 fillers, of which 20 were items from an independent experiment looking at the interpretation of possessive pronouns in sentences like (24).

- (24) Thomas finally persuaded Peter to publish his novel.

Each sentence was followed by a paraphrase with a gap as in (25). Participants' task was to complete the gap.

- (25) The policeman called the postman before he tied his shoelaces.
_____ tied his shoelaces.

Procedure: Three practice items were presented at the beginning of the experiment in order to familiarize participants with the task. Experimental items and fillers were then presented in individually randomized order. The experiments took about 20 minutes to complete. We provide the instructions in Appendix B.

Results: Answers were coded as 0 for Subject choices and 1 for Object choices. Completions other than Subject or Object choices were coded as Missing Values. Less than 3% of the data were affected by this procedure for both languages. We then computed a logistic mixed-effects model using the *lme4* package in R (Bates, Maechler, Bolker & Walker 2015) with Antecedent Choice (Object, Subject) as dependent variable, Language (French, English) and Connector (*before/avant*, *after/après*) as mean centered predictor variables, as well as random intercepts and slopes for Connector for participants and items, and the by-item random slope for Language as well as all possible interactions.

¹⁰ Note that the English materials are not always exact translations of the French materials. In French, we only used gender marked nouns to make sure that both antecedents have the same gender. This is not always possible in English. *Teacher* is for example unmarked for gender in English, while it is marked in French (fem: *la maîtresse*, masc: *le maître*). To make sure that co-occurring antecedents in our English items were considered approximately equally male or female respectively, we used the norms for the perception of role nouns provided in Misersky et al. (2014).

Before testing our experimental predictors, we tested for a potential systematic bias of the order of the nouns in our experiments, possibly due to semantic bias. We used antecedent choice as the dependent variable and Order as the predictor as well as random intercepts and slopes for participants and items. No systematic effect could be established (Estimate = -0.007785 , Std. Error = 0.202414 , $z = -0.038$, $p = 0.969$).¹¹

With respect to our experimental predictors, a significant effect of Language was found, with more Object choices for French than for English (Estimate = 1.6234 , Std. Error = 0.3259 , $z = 4.981$, $p < .001$) as well as a significant interaction of Language and Connector (Estimate = 1.7103 , Std. Error = 0.3122 , $z = 5.479$, $p < .001$). No other effects were significant. Subset models for each language separately showed that there were significantly more object choices with the connector *after* in English than with *before* (Estimate = -0.8123 , Std. Error = 0.2664 , $z = -3.050$, $p < 0.003$). The opposite pattern was found for French with fewer Object antecedent choices for *après* than for *avant* (Estimate = 0.9314 , Std. Error = 0.2552 , $z = 3.650$, $p < 0.001$).

Figure 6 shows the proportion of object choices for the connectors *after/après* and *before/avant* in English and French, and the model predictions are compared with the experimental results in Table 13.

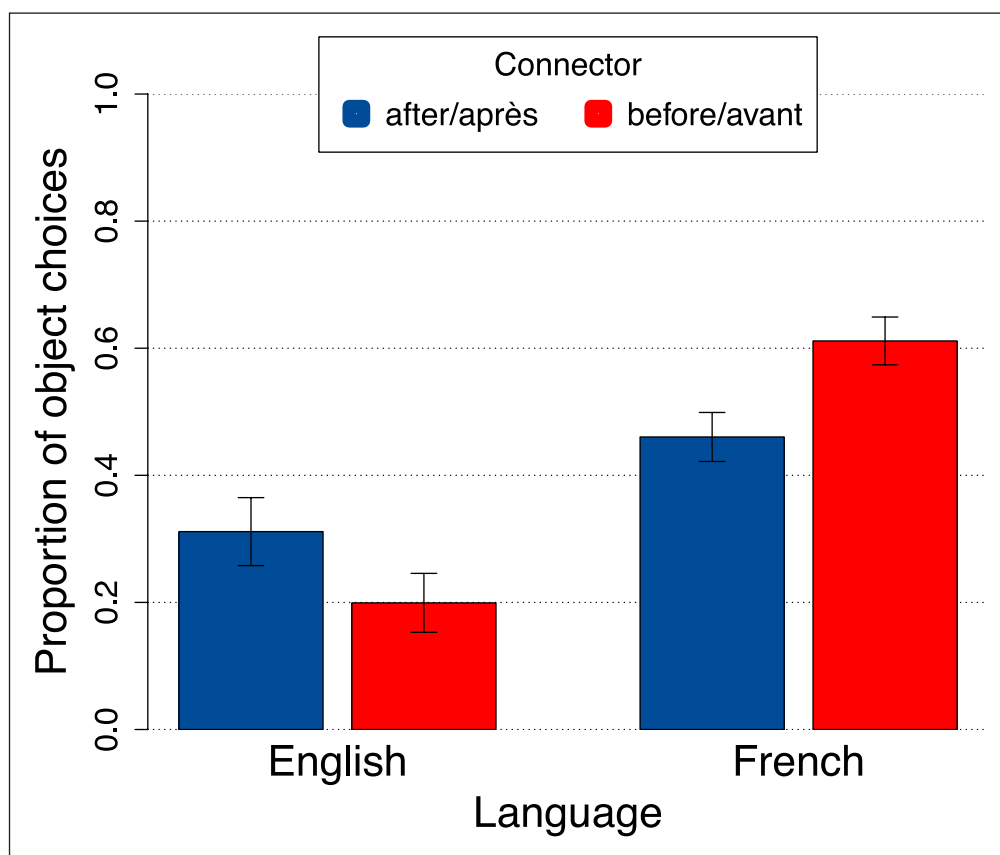


Figure 6 Proportion of Object choices for the connectors *after/après* and *before/avant* for English and French.

	Model predictions	Experimental results
English <i>after</i>	0.32	0.31
English <i>before</i>	0.23	0.20
French <i>après</i>	0.54	0.46
French <i>avant</i>	0.51	0.62

Table 13 Comparison of model predictions and experimental results.

¹¹ Although the lack of a significant effect of Order already shows that there can be no systematic effect of plausibility, following the suggestion of one of our anonymous reviewers, we further checked for any variation in plausibility (i.e., do participants prefer one noun over the other for any given pair of nouns, independent of their grammatical role in the matrix clause). We found that there was only slight variation (a 68/32 % bias in the worst case) which in any case was controlled by our Order variation in the experiment.

Discussion: The experimental data show the predicted language difference with more object antecedent choices for French than for English. The increased number of object choices for *after* for English participants was also predicted. The model does not, however, predict much difference in antecedent choices between *avant* and *après* in French. While the model predicts that *avant* and *après* should behave roughly the same since the ratio of the finite with respect to the non-finite construction is highly similar for both connectors (see [Tables 4](#) and [5](#)), the experimental results show a significantly greater object preference with *avant* than with *après*.

A possible way to explain this unexpected difference is the general frequency of *before/avant* and *after/après* in English and French. In English spoken corpora, *after* and *before* are roughly equally distributed (57% *before*, 43% *after*; see [Figure 3](#)). In spoken French, the ratio of *avant* to *après* is nearly 4 to 1 (78% *avant*, 22% *après*). It is possible that not only the relative frequency of the finite and non-finite constructions but also the general frequency of the connector will have to be taken into account in a more sophisticated model. Likewise, it is possible that the choice of indicative or subjunctive mood is made differently with *avant* and *après*, and that this difference has some effect on pronoun resolution. Although we think that the alternatives considered in our model are the closest and most plausible ones, as one anonymous reviewer points out, it would further be interesting to investigate a broader pool of (potentially language- or connector-specific) alternative constructions. This might allow us to draw a more differentiated picture, as could integrating item-specific variation into the model instead of assuming a uniform prior. However, we leave a more detailed investigation of the differences between *avant* and *après* to future work.

6 Conclusions

This paper presented new corpus and experimental studies of cross-linguistic variation in pronoun resolution. We tested the hypotheses that pronoun resolution preferences can be derived from Gricean reasoning, and that the relative frequency of a construction has an important effect on how it interacts with this alternative-based reasoning. In order to test these hypotheses in an explicit manner, we built Rational Speech Act models of pronoun resolution in English, French and German, where we incorporated alternative relative frequency into the models as costs.

The results in our paper generally support these proposals: we showed through a series of corpus studies of English and French that the finite constructions were much more frequent than the non-finite alternatives in English; however, the relative frequencies of (non)finite constructions in French were reversed. Consequently, our RSA models predicted more object interpretations in French than in English, and we showed that these predictions were borne out in new experimental studies of pronoun resolution in these two languages. This being said, we found differences in more fine-grained predictions of the models for the different connectors: while English *after* was correctly predicted to have more object interpretations than English *before*, our models predicted no significant difference between *avant* and *après*. In fact, *avant* had a significantly higher rate of object interpretations than *après* in the experimental data. Based on these results, we therefore conclude that Gricean reasoning + alternative frequency is not the full story when it comes to pronoun resolution: such models – at least as parametrized here – leave the difference between *avant* and *après* unexplained. Our paper thus also highlights how formal game-theoretic models, such as the RSA framework, can help us test high-level theoretical predictions about the relation between use, reasoning and interpretation and pinpoint key data for further investigation.

Abbreviations

COCA = Corpus of Contemporary American English, CQL = Corpus Query Language, ESLO = Enquête sociolinguistique à Orléans, INF = infinitive, NP = noun phrase, RSA = Rational Speech Act.

Additional file

The additional file for this article can be found as follows:

- **Appendices:** Details of the corpus analyses, experimental instructions, statistical formulas and details of the follow-up study. <https://osf.io/e4yfq/>

Ethics and consent

The experiments run for this paper followed standard ethics guidelines concerning consent and voluntary participation. All participants were informed about the experimental procedure and future data processing before providing their consent. Participation was voluntary and they could stop participating at any point without any data being stored.

Acknowledgements

We wish to thank our three anonymous reviewers for their useful and constructive comments, which significantly strengthened the final version of this work.

Funding information

This work has been partially funded by the Labex EFL (ANR/CGI 10-LABX-0083), which contributes to the IdEx Université de Paris (ANR-18-IDEX-0001).

Competing interests

The authors have no competing interests to declare.

Author affiliations

Miriam Schulz  orcid.org/0000-0002-9708-7936

Department of Language Science and Technology (LST), Universität des Saarlandes, 66123 Saarbrücken, DE

Heather Burnett  orcid.org/0000-0002-7389-0679

Laboratoire de Linguistique Formelle, Université de Paris – CNRS, 5, rue Thomas Mann – F-75205 Paris Cedex 13, FR

Barbara Hemforth  orcid.org/0000-0001-8371-1323

Laboratoire de Linguistique Formelle, Université de Paris – CNRS, 5, rue Thomas Mann – F-75205 Paris Cedex 13, FR

References

- Anderson, John Robert. 1991. Is human cognition adaptive? *Behavioral and Brain Sciences* 14(3). 471–485. DOI: <https://doi.org/10.1017/S0140525X00070801>
- Anthony, Laurence. 2018. *AntConc (Version 3.4.3)[Computer Software]*. Tokyo, Japan: Waseda University. (<https://www.laurenceanthony.net/software>) (Accessed 30-08-2019).
- Arnold, Jennifer E. 2001. The effect of thematic roles on pronoun use and frequency of reference continuation. *Discourse Processes* 31(2). 137–162. DOI: https://doi.org/10.1207/S15326950DP3102_02
- ATILF (Analyse et traitement informatique de la langue française – UMR 7118) & CLLE (Cognition, Langue, Langages, Ergonomie – UMR 5263). 2020. *Corpus journalistique issu de l'Est Républicain [Corpus]*. ORTOLANG (Open Resources and TOols for LANGUAGE) (https://hdl.handle.net/11403/est_republicain/v4) (Accessed 30-07-2020).
- Atlas, Jay David & Stephen Levinson. 1981. It-clefts, informativeness and logical form: Radical pragmatics (revised standard version). In Peter Cole (ed.), *Radical Pragmatics*, 1–61. New York: Academic Press.
- Bates, Douglas, Martin Maechler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). 1–48. DOI: <https://doi.org/10.18637/jss.v067.i01>
- Baumann, Peter, Lars Konieczny & Barbara Hemforth. 2011. Pragmatic expectations and coreference: How alternative constructions and referring expressions can serve as cues. In Laura Carlson, Christoph Hoelscher & Thomas F. Shipley (eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society*. Austin, TX: Cognitive Science Society. Retrieved from: <https://escholarship.org/uc/item/65d311dt>
- Baumann, Peter, Lars Konieczny & Barbara Hemforth. 2014. Conversational implicatures in anaphora resolution: Alternative constructions and referring expressions. In Barbara Hemforth, Barbara Mertins & Cathrine Fabricius-Hansen (eds.), *Psycholinguistic Approaches to Meaning and Understanding across Languages, Studies in Theoretical Psycholinguistics* 44, 197–212. Dordrecht: Springer. DOI: https://doi.org/10.1007/978-3-319-05675-3_8

- Canut, Cécile & Gudrun Ledegen. 1998. Après que... ou la fluctuation des modes en français parlé. *Langage et société* 85. 25–53. DOI: <https://doi.org/10.3406/lsoc.1998.2823>
- Carminati, Maria Nella. 2002. *The processing of Italian subject pronouns*. Amherst, MA: University of Massachusetts dissertation.
- Chierchia, Gennaro. 2004. Scalar implicatures, polarity phenomena, and the syntax/pragmatics interface. In Adriana Belletti (ed.), *Structures and Beyond* 3. 39–103. Oxford, UK: Oxford University Press.
- Colonna, Saveria, Sarah Schimke & Barbara Hemforth. 2012. Information structure effects on anaphora resolution in German and French: A cross-linguistic study of pronoun resolution. *Linguistics* 50(5). 991–1013. DOI: <https://doi.org/10.1515/ling-2012-0031>
- Colonna, Saveria, Sarah Schimke & Barbara Hemforth. 2014. Information structure and pronoun resolution in German and French: Evidence from the visual-world paradigm. In Barbara Hemforth, Barbara Mertins & Cathrine Fabricius-Hansen (eds.), *Psycholinguistic Approaches to Meaning and Understanding across Languages, Studies in Theoretical Psycholinguistics* 44, 175–195. Dordrecht: Springer. DOI: https://doi.org/10.1007/978-3-319-05675-3_7
- Cuetos, Fernando & Don C. Mitchell. 1988. Cross-linguistic differences in parsing: Restrictions on the use of the Late Closure strategy in Spanish. *Cognition* 30(1). 73–105. DOI: [https://doi.org/10.1016/0010-0277\(88\)90004-2](https://doi.org/10.1016/0010-0277(88)90004-2)
- Davies, Mark. 2008. *The Corpus of Contemporary American English: 450 million words, 1990–present*. (<http://corpus.byu.edu/coca/>) (Accessed 30-08-2019).
- De la Fuente, Israel & Barbara Hemforth. 2013. Effects of clefting and left-dislocation on subject and object pronoun resolution in Spanish. In Jennifer Cabrelli Amaro, Gillian Lord, Ana de Prada Pérez & Jessi Elana Aaron (eds.), *Selected Proceedings of the 16th Hispanic Linguistics Symposium*, 27–45. Florida: Cascadilla Proceedings Project.
- De la Fuente, Israel, Barbara Hemforth, Saveria Colonna & Sarah Schimke. 2016. The role of syntax, semantics, and pragmatics in pronoun resolution: A cross-linguistic overview. In Anke Holler and Katja Suckow (eds.), *Experimental Perspectives on Anaphora Resolution*, 11–32. Berlin, Boston: De Gruyter. DOI: <https://doi.org/10.1515/9783110464108-003>
- Fedele, Emily & Elsi Kaiser. 2014. Looking back and looking forward: Anaphora and cataphora in Italian. *University of Pennsylvania Working Papers in Linguistics* 20(1). (<https://repository.upenn.edu/pwpl/vol20/iss1/10>) (Accessed 30-08-2019).
- Fox, Danny & Roni Katzir. 2011. On the characterization of alternatives. *Natural Language Semantics* 19(1). 87–107. DOI: <https://doi.org/10.1007/s11050-010-9065-3>
- Frank, Michael C. & Noah D. Goodman. 2012. Predicting pragmatic reasoning in language games. *Science* 336(6084). 998. DOI: <https://doi.org/10.1126/science.1218633>
- Franke, Michael. 2009. *Signal to act: Game theory in pragmatics*. Amsterdam: University of Amsterdam dissertation.
- Frazier, Lyn & Charles Clifton. 1996. *Construal*. Cambridge, MA: MIT Press.
- Gazdar, Gerald. 1979. *Pragmatics: Presupposition, implicature, and logical form*. New York: Academic Press.
- Gernsbacher, Morton Ann & David J. Hargreaves. 1988. Accessing sentence participants: The advantage of first mention. *Journal of Memory and Language* 27. 699–717. DOI: [https://doi.org/10.1016/0749-596X\(88\)90016-2](https://doi.org/10.1016/0749-596X(88)90016-2)
- Goodman, Noah D. & Andreas Stuhlmüller. 2013. Knowledge and implicature: Modeling language understanding as social cognition. *Topics in Cognitive Science* 5. 173–184. DOI: <https://doi.org/10.1111/tops.12007>
- Goodman, Noah D. & Andreas Stuhlmüller. 2015. *The Design and Implementation of Probabilistic Programming Languages*. (<http://dippl.org/>) (Accessed 30-08-2019).
- Grice, Herbert Paul. 1975. Logic and conversation. In Peter Cole & Jerry L. Morgan (eds.), *Syntax and Semantics* 3. 41–58. New York: Academic Press.
- Hemforth, Barbara, Antoine Hédier, Theodor Cucu, Yair Haendler, Lisa Hemforth, Hugo Techer, Joanna Dos Reis & Doriane Gras. 2020. Verbal mismatch in French Right-Node Raising: Speeded grammaticality judgments and EEGs. *Talk at Experimental and Corpus-based Approaches to Ellipsis (ECBAE)*, Florence, Italy. (<http://drehu.linguist.univ-paris-diderot.fr/ecbae-2020/fichiers/abstracts/hemforth.pdf>) (Accessed 03-03-2021).
- Hemforth, Barbara, Lars Konieczny, Christoph Scheepers, Saveria Colonna, Sarah Schimke, Peter Baumann & Joël Pynte. 2010. Language specific preferences in anaphor resolution: Exposure or gricean maxims? In Stellan Ohlsson & Richard Catrambone (eds.), *Proceedings of the 32nd Annual Conference of the Cognitive Science Society*, 2218–2223. Portland: Cognitive Science Society.
- Hemforth, Barbara, Saveria Colonna, Joël Pynte & Lars Konieczny. 2004. Pronoun resolution across languages. *10th Annual Conference on Architectures and Mechanisms for Language Processing (AMLaP)*. Aix en Provence. (<http://www2.lpl-aix.fr/~fulltext/2037.pdf>) (Accessed 30-08-2019).
- Horn, Laurence R. 1972. *On the semantic properties of logical operators in English*. Los Angeles: UCLA dissertation.

- Horn, Laurence R. 1984. Toward a new taxonomy for pragmatic inference: Q-based and R-based implicature. In Deborah Schiffrin (ed.), *Meaning, form, and use in context: Linguistic applications*, 11–42.
- Horn, Laurence R. 1989. *A natural history of negation*. Chicago: Chicago University Press.
- Järvikivi, Juhani, Roger P. G. van Gompel, Jukka Hyönä & Raymond Bertram. 2005. Ambiguous pronoun resolution: Contrasting the first-mention and subject preference accounts. *Psychological Science* 16. 260–264. DOI: <https://doi.org/10.1111/j.0956-7976.2005.01525.x>
- Kastronic, Laura. 2016. *Going back to the source: A comparative analysis of the expression of necessity in Hexagonal and Quebec French*. Ottawa: University of Ottawa dissertation.
- Katzir, Roni. 2007. Structurally-defined alternatives. *Linguistics and Philosophy* 30(6). 669–690. DOI: <https://doi.org/10.1007/s10988-008-9029-y>
- Kehler, Andrew & Hannah Rohde. 2019. Prominence and coherence in a Bayesian theory of pronoun interpretation. *Journal of Pragmatics* 154. 63–78. DOI: <https://doi.org/10.1016/j.pragma.2018.04.006>
- Koehn, Philipp. 2005. Europarl: A Parallel Corpus for Statistical Machine Translation. *MT Summit*. (<http://homepages.inf.ed.ac.uk/pkoehn/publications/europarl-mtsummit05.pdf>) (Accessed 30-08-2019).
- Kroch, Anthony. 1972. Lexical and inferred meanings for some time adverbs. *Quarterly Progress Report of the Research Laboratory of Electronics* 104. 260–267. Cambridge, MA: MIT.
- Legendre, Géraldine. 1993. Antipassive with French psych-verbs. In Erin Dunca, Donka Farkas & Philip Spaelti (eds.), *Proceedings of the Twelfth West Coast Conference on Formal Linguistics*, 373–388. Stanford, CA: CSLI Publications.
- Levinson, Stephen C. 2000. *Presumptive meanings: The theory of generalized conversational implicature*. Cambridge, MA: MIT press. DOI: <https://doi.org/10.7551/mitpress/5526.001.0001>
- Lewis, David K. 1969. *Convention: A philosophical study*. Cambridge, MA: Harvard University Press.
- Luce, R. Duncan. 1959. *Individual choice behavior: A theoretical analysis*. New York: John Wiley & Sons.
- Mayol, Laia & Robin Clark. 2010. Pronouns in Catalan: Games of partial information and the use of linguistic resources. *Journal of Pragmatics* 42(3). 781–799. DOI: <https://doi.org/10.1016/j.pragma.2009.07.004>
- Misersky, Julia, Pascal M. Gygax, Paolo Canal, Ute Gabriel, Alan Garnham, Friederike Braun, Tania Chiarini, Kjellrun Englund, Adriana Hanulikova, Anton Öttl, Jana Valdova, Lisa von Stockhausen & Sabine Sczesny. 2014. Norms on the gender perception of role nouns in Czech, English, French, German, Italian, Norwegian, and Slovak. *Behavior research methods* 46(3). 841–871. DOI: <https://doi.org/10.3758/s13428-013-0409-z>
- Osborne, Martin J. & Ariel Rubinstein. 1994. *A course in game theory*. Cambridge, MA: MIT Press.
- Poplack, Shana, Allison Leales & Nathalie Dion. 2013. The evolving grammar of the French subjunctive. *International Journal of Latin and Romance Linguistics* 25(1). 139–195. DOI: <https://doi.org/10.1515/probus-2013-0005>
- Qing, Ciyang & Michael Franke. 2015. Variations on a Bayesian theme: Comparing bayesian models of referential reasoning. In Henk Zeevat & Hans-Christian Schmitz (eds.), *Bayesian Natural Language Semantics and Pragmatics. Language, Cognition, and Mind* 2. 201–220. Cham: Springer. DOI: https://doi.org/10.1007/978-3-319-17064-0_9
- Schimke, Sarah, Israel De la Fuente, Barbara Hemforth & Saveria Colonna. 2018. First language influence on second language offline and online ambiguous pronoun resolution. *Language Learning* 68(3). 744–779. DOI: <https://doi.org/10.1111/lang.12293>
- Schulz, Miriam. 2018. *Investigating the impact of alternative constructions on cross-linguistic preferences in pronoun resolution: A corpus study and Rational Speech Act model*. Paris: Université Paris 7 – Diderot unpublished master's thesis.
- Scontras, Gregory, Michael Henry Tessler & Michael Franke. 2017. *Probabilistic language understanding: An introduction to the Rational Speech Act framework*. (<https://www.problang.org>) (Accessed 30-08-2019).
- Seddah, Djamel, Marie Candito, Benoît Crabbé & Henestroza Anguiano Enrique. 2012. Ubiquitous usage of a French large corpus: Processing the Est Republicain corpus. In Nicoletta Calzolari, Khalid Choukri, Thierry Declerck, Mehmet Uğur Doğan, Bente Maegaard, Joseph Mariani, Jan Odijk & Stelios Piperidis (eds.), *Proceedings of the Eighth International Conference on Language Resources and Evaluation*, 3249–3254. Istanbul, Turkey: European Language Resources Association (ELRA).
- Shannon, Claude. 1948. A mathematical theory of communications. *Bell Systems Technical Journal* 27(4). 623–656. DOI: <https://doi.org/10.1002/j.1538-7305.1948.tb00917.x>
- Sutton, Richard S. & Andrew G. Barto. 1998. *Reinforcement learning: An introduction*. Cambridge, MA: MIT Press. DOI: <https://doi.org/10.1109/TNN.1998.712192>

TO CITE THIS ARTICLE:

Schulz, Miriam, Heather Burnett and Barbara Hemforth. 2021. Corpus, experimental and modeling investigations of cross-linguistic differences in pronoun resolution preferences. *Glossa: a journal of general linguistics* 6(1): 66. 1–21. DOI: <https://doi.org/10.5334/gjgl.1142>

Submitted: 21 October 2019

Accepted: 10 February 2021

Published: 14 May 2021

COPYRIGHT:

© 2021 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.

Glossa: a journal of general linguistics is a peer-reviewed open access journal published by Ubiquity Press.