

Linguistic evidence across languages

Barbara Hemforth

Linguistic Evidence 2018

Tübingen

Overview

- Some things about crosslinguistic studies that I want to get of my chest
- Pronouns
 - Baseline differences across languages: alternative constructions
 - Baseline differences across languages: alternative pronominal forms
 - Adaptation to global and local statistics

How language comparison is typically done in psycholinguistics

Hey, people found a subject relative preference for English? Does that work for my language/ language X as well? Let's set up some materials and check.

- The processing of subject and object relative clauses in Spanish: An eye-tracking study
- Animacy effects in Chinese relative clause processing
- An ERP study of the processing of subject and object relative clauses in Japanese
- **Recommendation:**
 - **If same effects as elsewhere: Important replication and necessary if the tested constraint is assumed to be language general (but is this for the same reasons?)**
 - **If different effects: Run a comparative control study**
 - But be really careful to get the translations right

Variation of preferences within languages

- **Not all object relatives are hard in languages that prefer subject relatives**
- **Not all relatives attach high in so called high attachment languages**
- **Not all pronouns take subject antecedents in a language showing a general preference for subjects**

Variation of preferences within languages

Solution:

- Take care of all the factors that we know of to be as close as possible to some experiment that has been published for another language

But still:

- There may be factors we don't know of yet.

The best solution:

- Run parallel studies with translation equivalent materials and the same experimental paradigm
 - Talk to a linguist for the translations
- If you can replicate preferences for some other language and still find something different in the language you are interested in you may be fine.

Is there variation in syntactic
island effects?

Ted's talk tomorrow!

Variation in pronoun resolution across languages and varieties

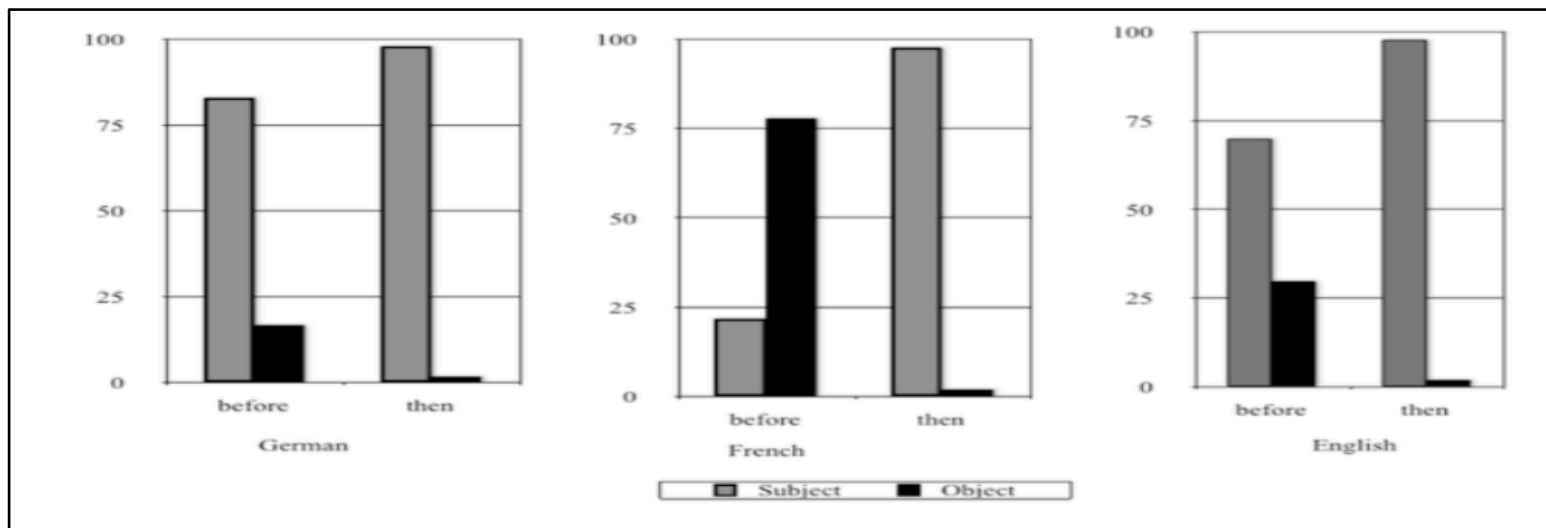
Collaborations

- Peter Baumann (Northwestern University)
- Heather Burnett (Paris Diderot)
- Saveria Colonna (Université Paris 8)
- Eduardo Correa Soares (Paris Diderot)
- Eunice Fernandes (Paris Diderot)
- Israel de la Fuente (Univ. Lille)
- Lars Konieczny (Universität Freiburg)
- Paula Luegi (Univ. Lisbon)
- Joël Pynte (Paris Descartes)
- Christoph Scheepers (University of Glasgow)
- Sarah Schimke (Universität Münster)

Baseline preferences across languages

- a. *The postman met the street-sweeper before he went home.*
- b. *The postman met the street-sweeper. Then he went home.*

The _____ went home.



Baseline preferences across languages

- In *before* (intra-sentential) contexts (offline and online):
 - Subject preference in English and German
 - Object preference in French
- In *then* (inter-sentential) contexts:
 - Subject preference in all three languages

Baseline preferences across languages

- Why does French differ from English and German?
- Availability of a (nearly) unambiguous construction in French:

Le facteur a rencontré le balayeur avant de ramasser les lettres.

‘The postman met the street-sweeper before picking up the letters’

- French participants might be following a Gricean Principle of Manner (‘avoid ambiguity’) that results in the observed division of labor
- While this alternative construction does not exist in German, it does in English (*before he* vs. *before –ing*)

Baseline preferences across languages

- Why does English pattern with German and not with French?
- *Europarl* corpus analysis:

	FRENCH	ENGLISH
Subordinate clause	1917	7460
Infinitive/gerundive construction	3043	1725
Ratio	1.58	0.23

How to explain crosslinguistic variation in pronoun resolution?

- Burnett & Hemforth (2017a,b) show how syntactic differences can give rise to pragmatic differences using **signaling game** models (Lewis 1969), particularly IBR/RSA instantiations (Franke 2009, Frank & Goodman 2012).

A common RSA architecture for language differences based on alternative constructions

- We use the iterated RSA solution concept for all three games. $\langle S, L, \{pS, pO\}, M, Pr \rangle$:
- S is the speaker; L is the listener.
- Propositions under consideration:
 - The individual denoted by the subject went home (pS) and
 - The individual denoted by the object went home. (pO)

Messages

- We assume that the three languages differ only in the properties of their messages: the M component of the game.
- Following Arnold (2001), we assume that hearing a DP in subject position increases L's expectation that this DP will serve as a referent in the subsequent discourse.
- English and French differ in the relative frequency of the PRO form. The overt pronoun form is 4.32 times more frequent than the PRO form; whereas, the PRO form was found to be 1.58 times more frequent in French studies (Baumann et al. 2014).

The Model:

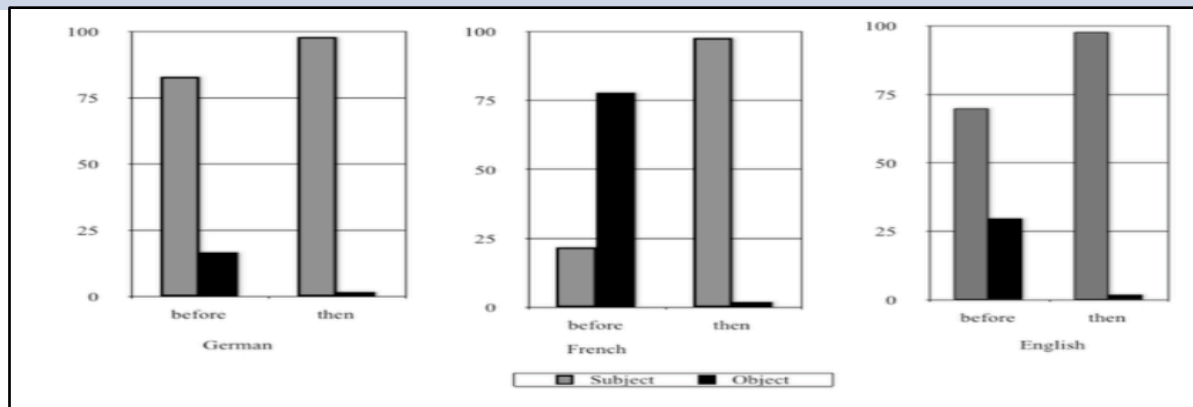
Less accessible constructions are more costly

Form (m)	German		English		French	
	$\llbracket m \rrbracket$	$\text{Cost}(m)$	$\llbracket m \rrbracket$	$\text{Cost}(m)$	$\llbracket m \rrbracket$	$\text{Cost}(m)$
Overt (<i>er/he/il</i>)	$\{p_S, p_O\}$	0	$\{p_S, p_O\}$	0	$\{p_S, p_O\}$	1.5
PRO			$\{p_S\}$	1	$\{p_S\}$	0

Model predictions and testing

LANGUAGE	BETWEEN SENTENCE (THEN)	WITHIN SENTENCE (BEFORE)
German	95%	95%
English	95%	72%
French	95%	15%

Table 1: RSA Model predictions ($\lambda = 1$)



Baseline preferences across languages: alternative pronominal forms

- Availability and frequency of alternative *forms of reference*
 - Null and overt subject pronouns in Spanish and European Portuguese
 - Stressed and unstressed pronouns in French
 - Weak and strong pronouns in German (“er” vs “der”)
- Carminati (2002) proposes a division of labor in the processing of null and overt pronouns in Italian
 - Null pronouns prefer the more salient subject antecedent
 - Overt pronouns prefer an antecedent in a different syntactic position (e.g. object)

Baseline preferences across languages: alternative pronominal forms

- Baumann, Konieczny, & Hemforth (2014): alternative referential forms and alternative constructions in European Portuguese

a. *O pintor viu o pescador, antes que (ele) abrisse a janela.*

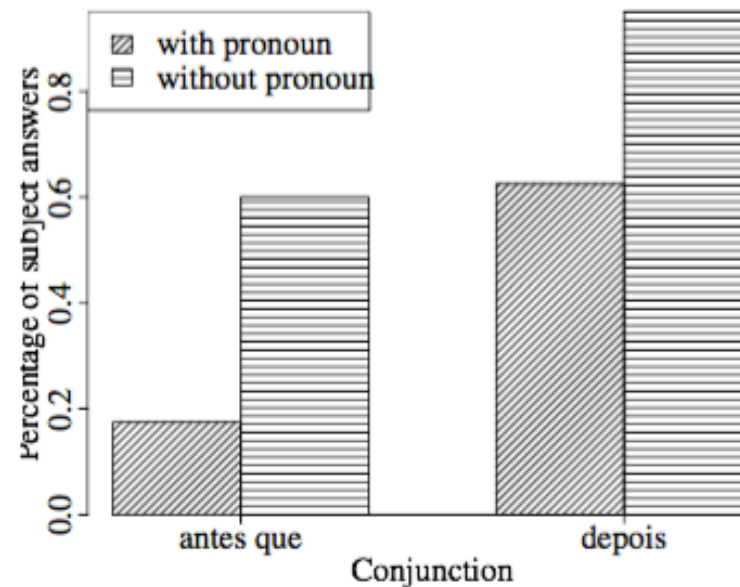
‘The painter saw the fisherman before (he) opened the window’

b. *O pintor viu o pescador. Depois (ele) abriu a janela.*

‘The painter saw the fisherman. After that (he) opened the window.’

O _____ abriu a janela.

Baseline preferences across languages: alternative pronominal forms



Baseline preferences across languages

- Self-paced reading study:

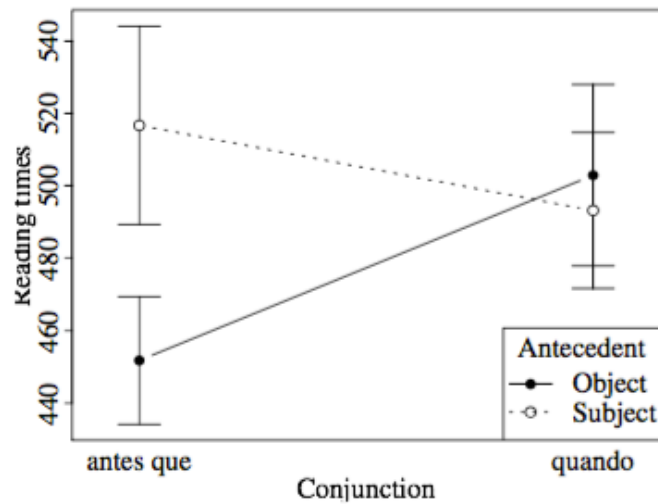
a. *O polícia encorajou a atriz, antes que **ele / ela** voltasse para casa.*

‘The policeman encouraged the actress before **he / she** went back home’

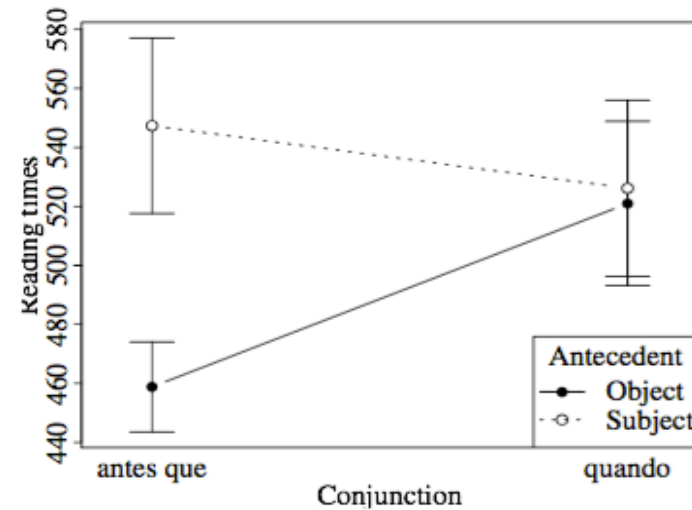
b. *O polícia encorajou a atriz, quand **ele / ela** voltou para casa.*

‘The policeman encouraged the actress when he / she went back home’

Baseline preferences across languages



Pronoun region



Spill-over region

- Faster RTs on pronoun and spill-over region when the pronoun referred to the object of the preceding clause after *antes que*
- No differences in RTs between the *quando* conditions

Baseline preferences across languages

- The results of these experiments suggest that the availability of both an alternative construction and a referential expression play a role in pronoun interpretation in Portuguese
- The results of the reading study indicate that, contra Carminati's predictions, the *overt* pronoun in Portuguese can refer to both the subject and the object antecedent
 - This pattern has also been observed in Spanish (e.g. de la Fuente, 2011)

Not only frequency in the
language counts

The local environment

- de la Fuente & Hemforth (2013): null vs. overt subject pronoun resolution in Spanish

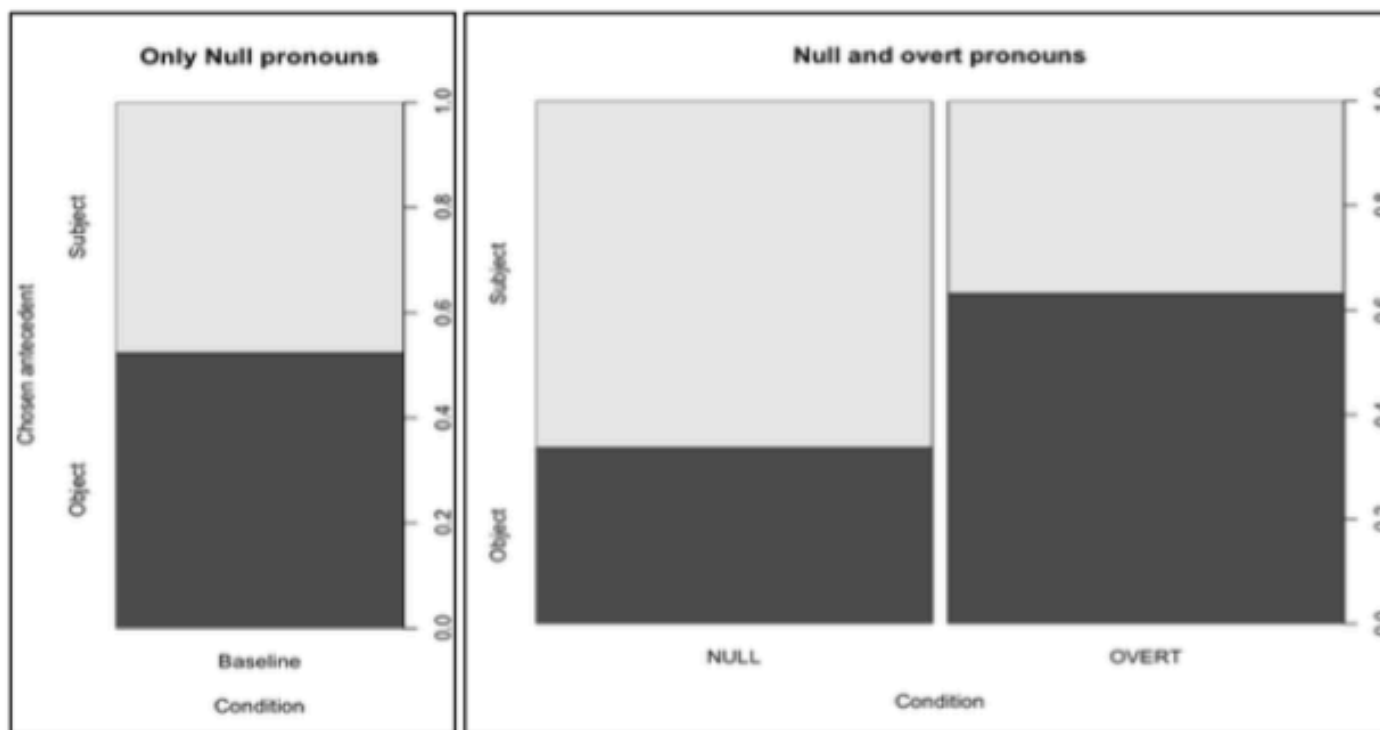
a. Juan llamó a Pedro cuando estaba en la oficina.

b. Juan llamó a Pedro cuando él estaba en la oficina.

‘Juan called Pedro when (he) was in the office’

- 2 questionnaire studies: only null pronouns vs. both null and overt pronouns

The local environment



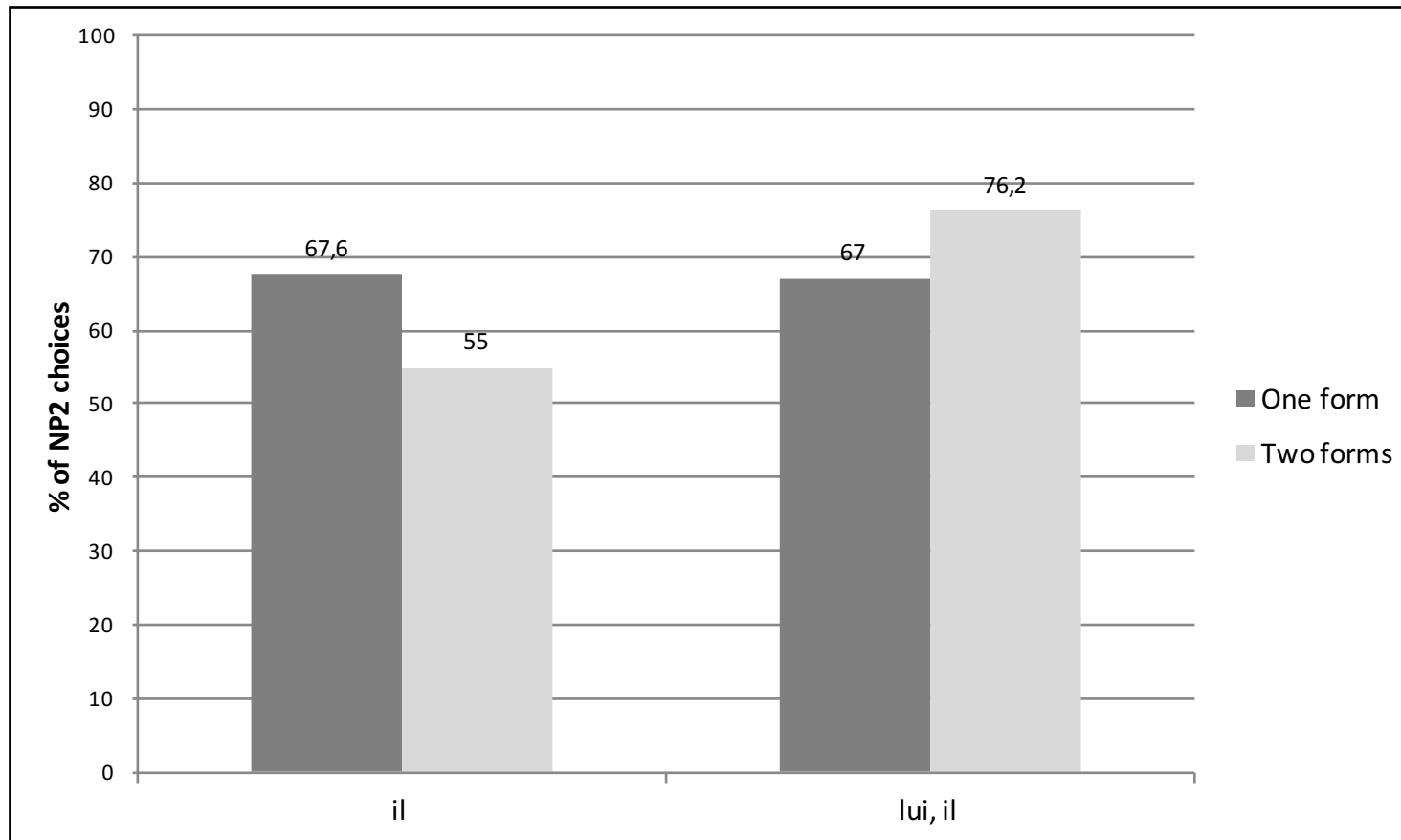
The local environment

- Colonna, Schimke, de la Fuente & Hemforth (2016): stressed vs. unstressed pronoun resolution in French

- (8) a. *Pierre a giflé Jean. Il était stagiaire.*
 b. *Pierre a giflé Jean. Lui, il était stagiaire.*
 ‘Pierre slapped Jean. He / HE was a trainee’

- 3 questionnaire studies: only unstressed pronouns, only stressed pronouns, both stressed and unstressed pronouns

Baseline preferences across languages



The local environment

- Different baseline (subject/object) antecedent preferences in the languages investigated
- The availability and frequency of alternative ambiguous constructions (e.g. subordinate vs. infinitival) and alternative forms of reference (e.g. null vs. overt, stressed vs. unstressed) can affect interpretation choices
 - Exposure effects both at the level of the particular grammar but also within the experiment
- Baseline preferences can also change depending on whether resolution occurs within or across sentence boundaries

Global and local frequency of alternative forms

- Cross-linguistic (Filiaci et al., 2013) as well as cross-dialectal (Luegi, 2012) differences have been attested: the Overt pronoun preference for the object is reduced (or absent) in Spanish and in Brazilian Portuguese (BP), when compared to Italian and European Portuguese (EP)
- Can the observed differences between BP vs. EP be explained in terms of the amount of exposure to different pronominal forms?
- In BP the occurrence of Overt pronouns is increasing and the use of Null pronouns is decreasing (e.g Duarte, 1995)

Global and local frequency of alternative forms

- Cross-linguistic (Filiaci et al., 2013) as well as cross-dialectal (Luegi, 2012) differences have been attested: the Overt pronoun preference for the object is reduced (or absent) in Spanish and in Brazilian Portuguese (BP), when compared to Italian and European Portuguese (EP)
- Can the observed differences between BP vs. EP be explained in terms of the amount of exposure to different pronominal forms?
- In BP the occurrence of Overt pronouns is increasing and the use of Null pronouns is decreasing (e.g Duarte, 1995)

Comparing European and Brazilian Portuguese

(Fernandes, Luegi, Correa Soares, de la Fuente, & Hemforth, in press)

- We tested Global (in the language, contrasting BP vs. EP) and Local (in the experimental context) exposure effects on Overt and Null pronoun resolution in Portuguese:
 - Global exposure effects should arise from contrasting BP and EP
 - Local exposure was manipulated in two exposure conditions where we varied the relative amount of Null and Overt forms
- If the processor is sensitive to pronoun frequency, we should observe:
 - An effect of Global exposure: the Overt pronoun will prefer the object antecedent in EP, and no preference in BP
 - An effect of Local exposure, increased exposure to Null forms should elicit in BP the division of labor observed in EP

What may we expect?

- Assumptions:
 - $P(\text{Subj}) = .70$; $P(\text{Obj}) = .30$
 - $P(\text{Subj} | \text{Null}) = .9$
 - $P(\text{Obj} | \text{Overt}) = .6$
- What we vary: $P(\text{Null})$ and $P(\text{Overt})$
- Different adaptations are possible:
 - They may keep the form-based likelihoods and adapt the expectations for subject and object antecedents
 - E.g. with 75% nulls, $P(\text{Subj}) = .775$; $P(\text{Obj}) = .225$
 - They may adapt $P(\text{Subj} | \text{Null})$ and keep the priors
 - $P(\text{Subj} | \text{Null}) = .8$, with 75 % nulls, $P(\text{Subj}) = .7$
 - They may adapt $P(\text{Obj} | \text{Overt}) = .99$, with 25% overt, $P(\text{Obj}) = .25$

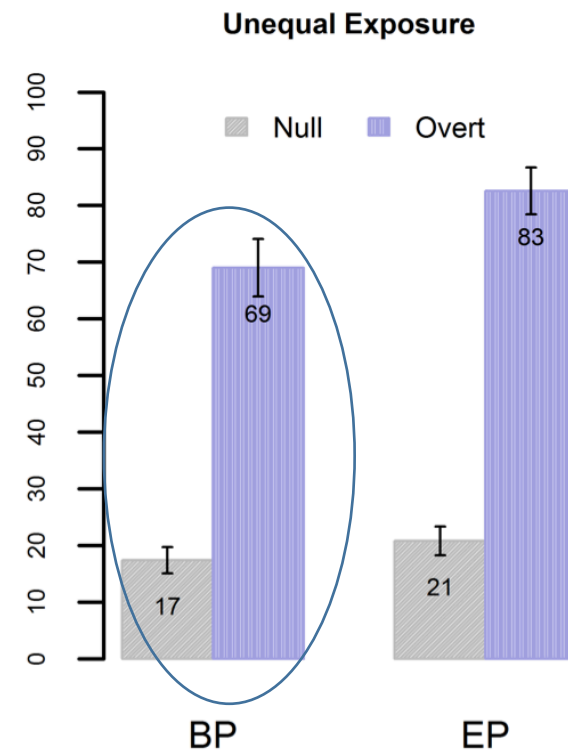
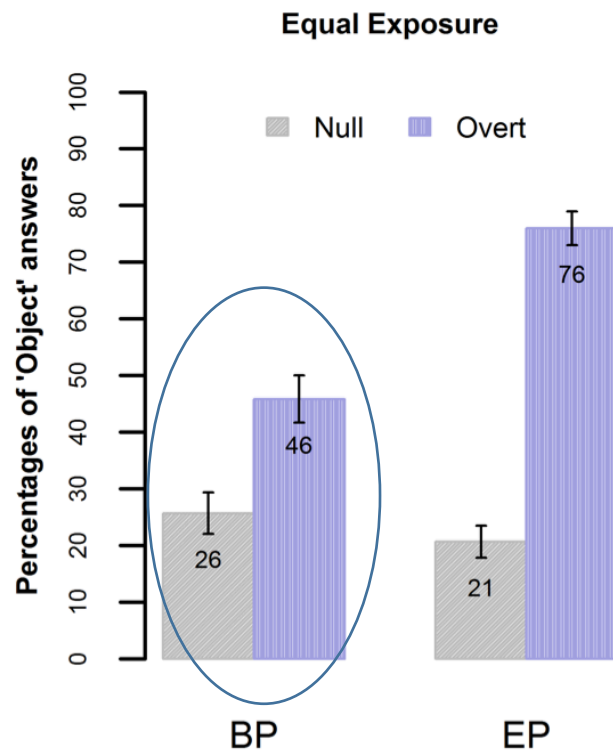
Questionnaire experiment

- 24 native speakers of EP (mean age 23.5 ± 7.1)
- 20 native speakers of BP (mean age 21.0 ± 3.6)
- 32 experimental sentences in 4 conditions (in each Variety) Exposure: 50%_(Null)/50%_(Overt) vs. 75%_(Null)/25%_(Overt)
- Pronoun: Null vs. Overt
 - a. O atleta consultou o ortopedista no hospital quando \emptyset regressou da viagem a Itá lia.
'The athlete consulted the orthopedist at the hospital when he returned from the journey to Italy.'
 - b. O atleta consultou o ortopedista no hospital quando ele regressou da viagem a Itá lia.
'The athlete consulted the orthopedist at the hospital when he returned from the journey to Italy.'
- 64 filler sentences
- Off-line internet-based questionnaire (IbexFarm)
- Interpretation question (e.g., Who returned from the journey?) with two possible answers (the athlete or the orthopedist).

Results

BP:
Equal exposure,
Object choices: .36

Unequal exposure,
Object choices: .30



Results

- Pronoun x Exposure interaction effect in BP (marginal in BP vs. EP): more object choices with Overt pronoun in different exposure condition
- Main effect of Pronoun across varieties: more object choices in Overt pronoun condition
- Main effect of Variety: more object choices in EP
Pronoun x Variety interaction: more object choices with Overt pronoun in EP

Global and local effects: Visual World experiment

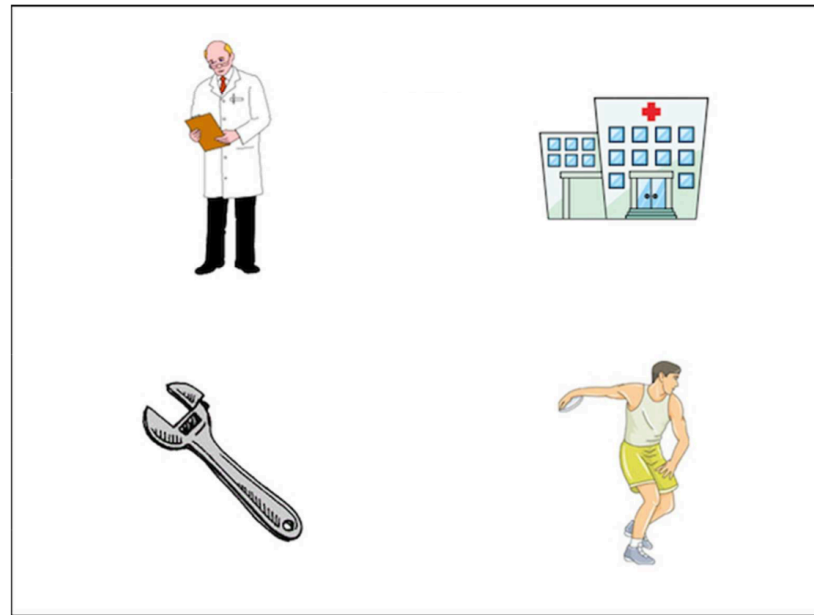


Figure 3. Visual context depicting the entities mentioned in the spoken target sentence The athlete consulted the orthopaedist at the hospital when he/ Ø returned from the journey to Italy: the pronoun antecedents (subject: ATHLETE and object: ORTHOPAEDIST), the location (HOSPITAL) and the distractor (SCREWDRIVER).

Results

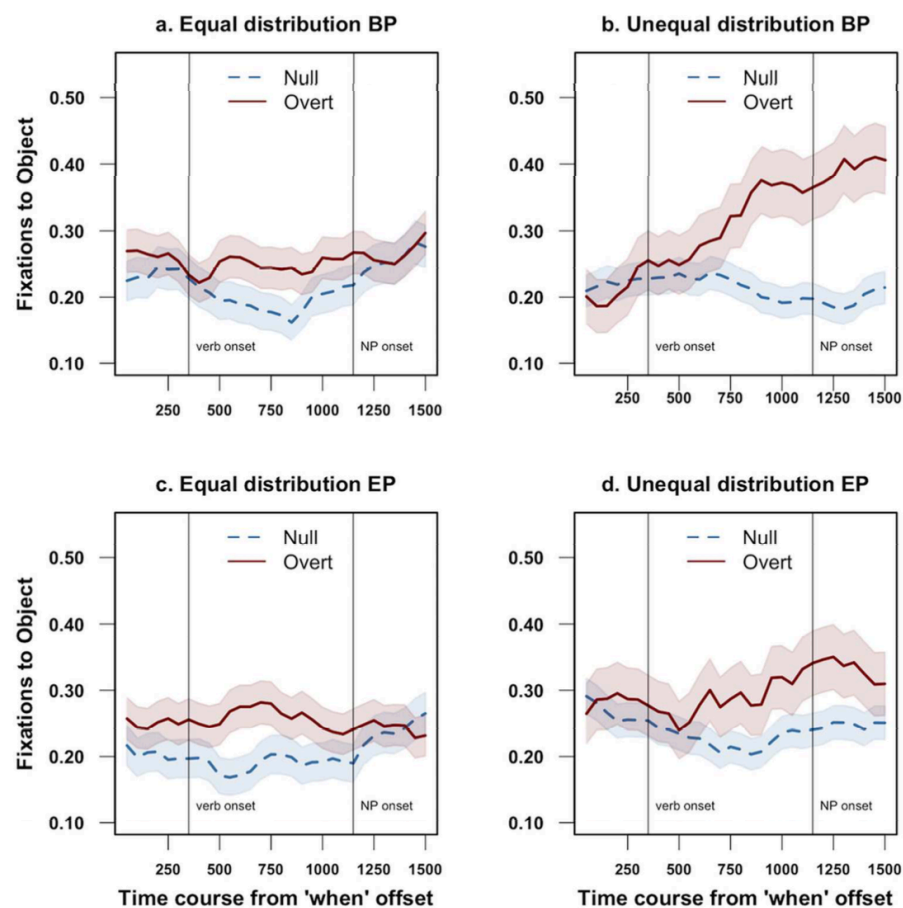


Figure 4. Experiment 2; Time-course of proportion of fixation to OBJ from the offset of when for BP (top row) and EP (bottom row) in the conditions crossing Pronoun (Null, dashed blue lines vs. Overt, solid red lines) and Distribution (Equal, left panels vs. Unequal, right panels). Shaded bands represent SEs around the observed means. The vertical lines indicate the mean onsets of the verb and of the following phrase.

Conclusions

- Exposure plays a role in Null and Overt pronoun resolution:
 - (1) In BP, contrary to what happens in EP, the division of labour is not observed for (more frequent) Overt pronouns, in an environment where the two forms are equally distributed
 - (2) However, this pattern emerges when exposure is skewed towards a higher relative frequency of Null pronouns
- These results extend previous evidence of Local exposure effects (de la Fuente & Hemforth, 2013, Colonna et al., 2016) and are in line with related evidence of effects of global availability of different constructions on pronoun resolution cross-linguistically (de la Fuente et al., 2016)
- Crucially, these results provide, for the first time, evidence of a Global and Local Exposure interaction and its effect on pronoun interpretation

What's going on here?

- Increasing accessibility of the null pronoun triggers division of labor effects in Brazilian Portuguese
 - General priors for subject antecedents are independent from the distribution of pronominal forms
 - Null pronouns have a strong subject bias across varieties
 - Increasing null pronouns forces overt pronouns to specialize

Conclusions

- What is adapted?
 - Prior expectations of object antecedents.
 - Only very slightly
 - Likelihood of $P(\text{Subj} | \text{Null})$?
 - Not much
 - Likelihood of $P(\text{Obj} | \text{Overt})$?
 - Very much so

Conclusions

1. Discourse expectations of what will be mentioned next (Subj or Obj) are fairly stable
 - Based on a larger amounts of data?
2. The preference of a null pronoun for a subject antecedent is very strong across languages
 - Are strong preferences harder to adapt?
3. Preferences for overt pronouns are adapted to keep 1 and 2 as constant as possible.

General conclusion

- Morpho-syntactic properties of languages as well as distributional properties of constructions and referential forms affect pronoun resolution.
- The accessibility of a form/construction determines its cost for the speaker. This cost is taken into account by the listener.
- Accessibility is determined by distributional properties in the language and/or in the immediate linguistic experience.

Danke!