Voice your opinion!
On syntactic islands in spoken English and the influence of voice.

Anne Mette Nyvad, Ken Ramshøj Christensen & Julie Maria Rohde

Dept. of English, AU

MINDS – Mind the structure!

https://tildeweb.au.dk/au572/minds.html

07.11.2023

Short background

- In the generative syntax literature, it is generally assumed that it is impossible to extract from adverbial clauses across languages:

  (1) ***Who** did Mary cry [after John hit __]?   (Huang 1982: 503)

- Huang (1982: 505) proposed the Condition on Extraction Domain (CED),
  - In a nutshell – states that subjects and modifiers are syntactic "islands" (= difficult/impossible to escape from). Hence, we should expect:
    - adjunct clause extraction to receive low acceptability ratings
    - and little or no variation in acceptability across constructions and languages.

Huang, Cheng-Teh James. 1982. *Logical Relations in Chinese and the Theory of Grammar*. Cambridge, MA: MIT PhD dissertation.

Short background

- The acceptability level appears to depend on a number of different factors:

- *Type of adverbial clause:*
  - In Norwegian and Swedish, movement out of adjunct clauses of condition (*if-clauses*) and time (*when-clauses*) are better than those of reason (*because-clauses*) (Bondevik et al. 2020; Müller 2017)
- *Type of dependency:*
  - Topicalization is more acceptable than *wh*-extraction (Kush et al. 2018, 2019)
- *Context:*
  - The addition of a supporting context increases the acceptability (Kush et al. 2019)

Bondevik, Ingrid, Dave Kush & Terje Lohndal. 2020. Variation in adjunct islands: The case of Norwegian. *Nordic Journal of Linguistics* 1–32. https://doi.org/10.1017/S0332586520000207.
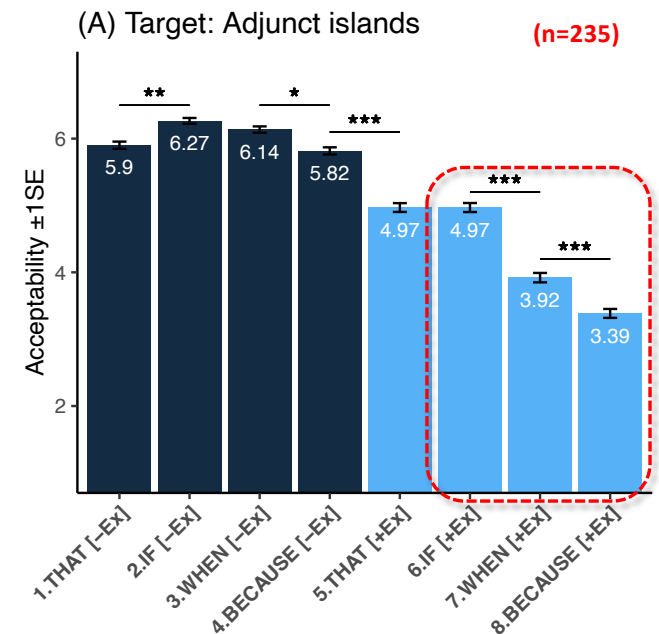Kush, Dave, Terje Lohndal & Jon Sprouse. 2018. Investigating variation in island effects: A Case Study of Norwegian Wh-Extraction. *Natural Language & Linguistic Theory* 36(3). 743–779. https://doi.org/10.1007/s11049-017-9390-z.
Kush, Dave, Terje Lohndal & Jon Sprouse. 2019. On the island sensitivity of topicalization in Norwegian: An experimental investigation. *Language* 95(3). 393–420. https://doi.org/10.1353/lan.2019.0051.
Müller, Christiane. 2017. Extraction from Adjunct Islands in Swedish. *Norsk Lingvistisk Tidsskrift* 35(1). 67–85.
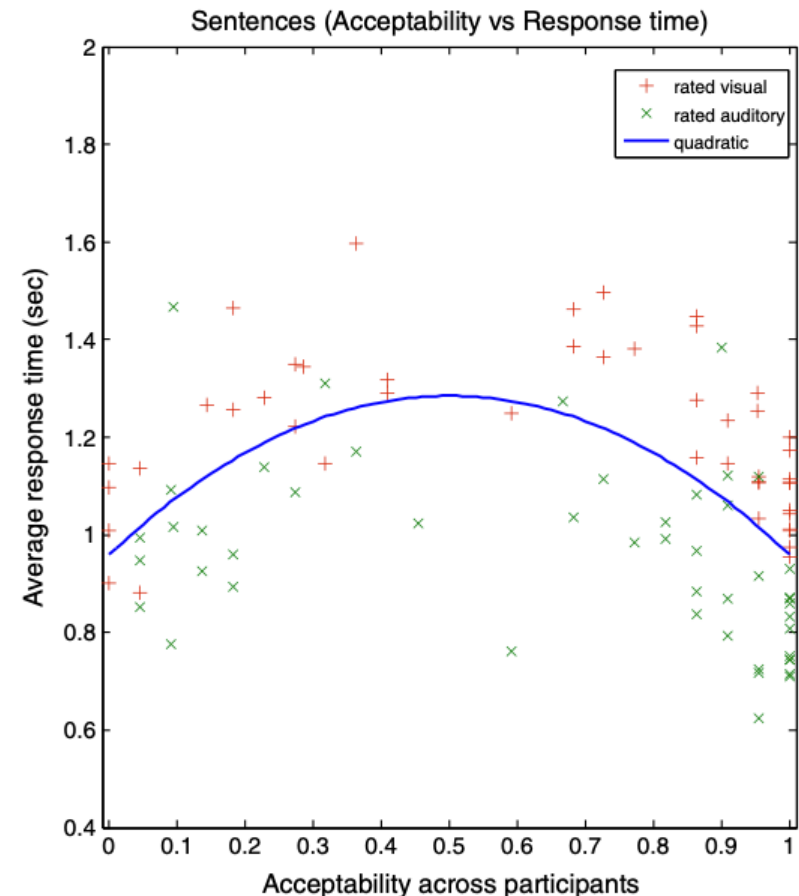
Results of Nyvad et al. (2022)

- **Heterogeneous pattern:**
  - Different acceptability levels *if > when > because*
  - Acceptability of *if*-clauses relatively high and not significant from *that*-clauses.
    - Conclusion: *If*-clauses are not strong islands in English.

- **In short, the acceptability of island extractions can be manipulated with linguistic factors.**

- **But what about VOICE?**

(A) Target: Adjunct islands (n=235)

Nyvad, Anne Mette, Christiane Müller, and Ken Ramshøj Christensen. 2022. "Too True to Be Good? The Non-Uniformity of Extraction from Adjunct Clauses in English." *Languages* 7 (4): 244. https://doi.org/10.3390/languages7040244.

Christensen & Wallentin (2010)

- **RT and acceptability**: Faster response with clearly acceptable or unacceptable items.
  - 'Uncertainty' slows you down.

- **RT and Modality** (written vs. spoken): Faster RT with spoken stimuli than with written stimuli
  - Significant main and interaction effects of modality on RT (all p>0.02)

- **Acceptability and Modality**: No main or interaction effects of modality (ANOVA, p>0.9)

Sentences (Acceptability vs Response time)

Christensen, Ken Ramshøj & Mikkel Wallentin. 2011. The locative alternation: Distinguishing linguistic processing cost from error signals in Broca's region. *NeuroImage* 56(3). 1622–1631. https://doi.org/10.1016/j.neuroimage.2011.02.081.

Present study: Stimuli

- We conducted an experiment using the following set of sentences, namely relativization out of adverbial clauses introduced by *if*, *when* and *because* (cf. Nyvad et al. 2022):

a. It's obvious that I was surprised [**that** she actually completed this exercise].   [-EX, -ISL]
b. It's obvious that I would be surprised [**if** she actually completed this exercise].   [-EX, +ISL]
c. It's obvious that I was surprised [**when/because** she actually completed this exercise].[-EX, +ISL]

e. This is <u>the exercise</u> that I was surprised [**that** she actually completed __].   [+EX, -ISL]
f. This is <u>the exercise</u> that I would be surprised [**if** she actually completed __].   **[+EX, +ISL]**
g. This is <u>the exercise</u> that I was surprised [**when/because** she actually completed __].   **[+EX, +ISL]**

Nyvad, Anne Mette, Christiane Müller, and Ken Ramshøj Christensen. 2022. "Too True to Be Good? The Non-Uniformity of Extraction from Adjunct Clauses in English." *Languages* 7 (4): 244. https://doi.org/10.3390/languages7040244.

Predictions

- Previous studies have used written stimuli. However, extraction from islands is (presumably) primarily a spoken language phenomenon. **Hence, we predict spoken to potentially ameliorate the island effect.**

- In order to explore the potential impact of difference between individual voices, we used a Female and a Male voice. AI generated speech with *IIElevenLabs.io*:
  - **Bella** - descriptors: Soft, narration
  - **Charlie** – descriptors: Casual, conversational

- Admittedly, we had no a priori prediction about any effect of F/M on acceptability/RT, but you never know…

Results

- Binary acceptability judgment task (OK / Not OK)
- Stimuli: 76 sentences in total (48 targets + 28 fillers). Latin square design. 2 lists. Pseudorandom assignment to list. Each participant judged 38 items.
- PsychoPy script running online on Pavlovia.org

- Participants: N = 48
  - All native speakers of English
  - Age: 17-72 years (mean = 44.2 years)
  - Gender: 24F, 24M [4 'other' excluded: too few data points]
  - List 1: 19, list 2: 29
  - Analysis excluded responses with RT ≥ 6 sec.

Results

- **Modality**: Same overall pattern as in our previous experiment with written stimuli.

(A) Target: Adjunct islands

Nyvad et al. (2022)

*** p < .001, ** p < .01, * p < .05

Acceptability (n=48)
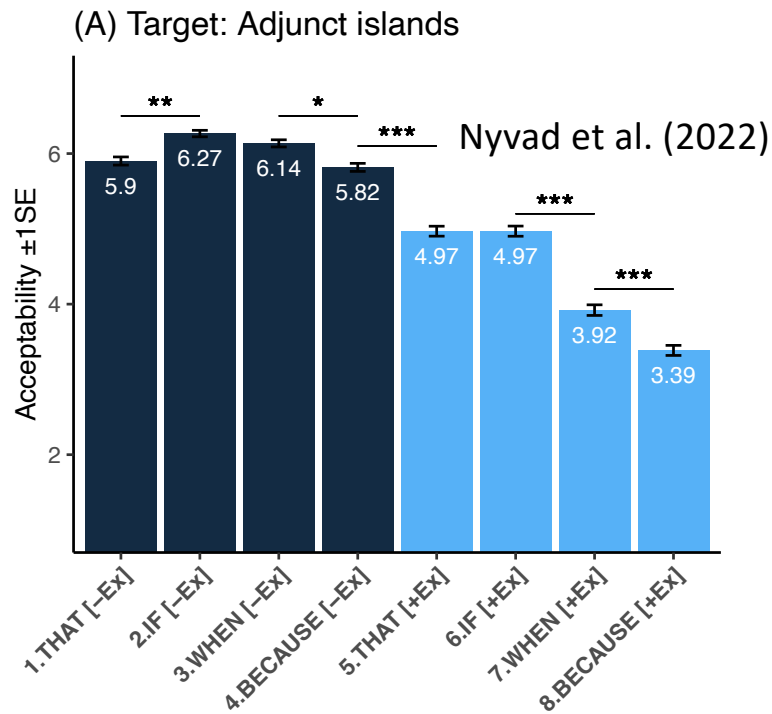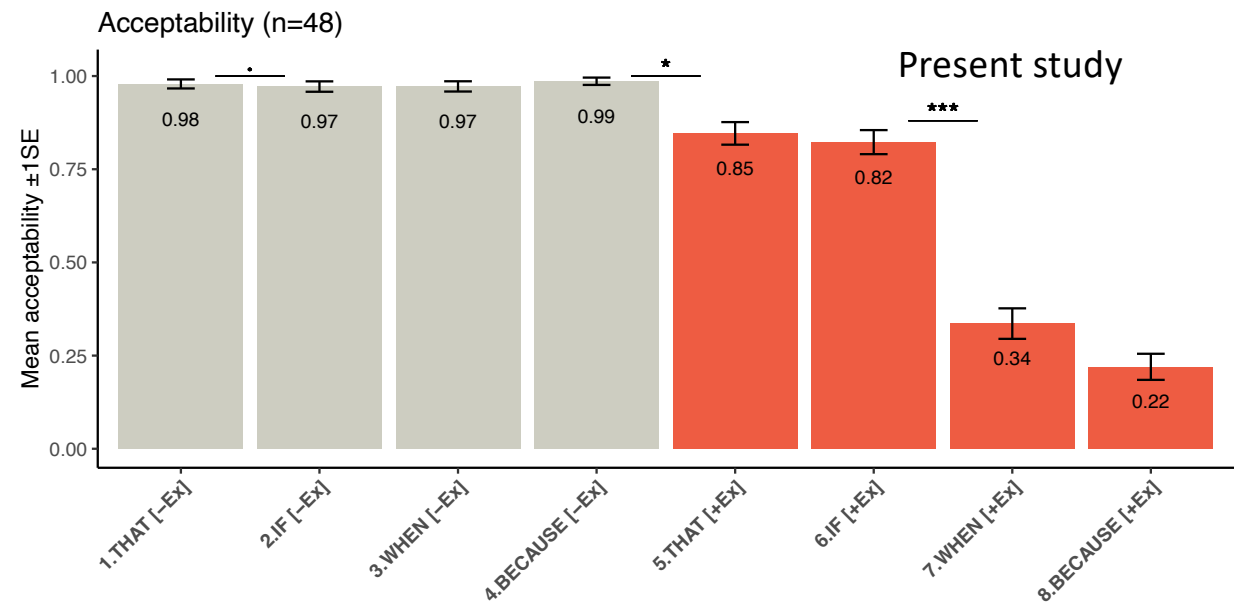
Present study

MODEL = glmer(ANSWER ~ TYPE + (1+TYPE|PARTICIPANT) + (1+TYPE+TRIAL|ITEM), data = DATA1, family = "binomial")

Modality & *z*-transformation

Comparing data from present study (spoken) and data from Nyvad et al. (2022) (written)

- A participant's z-transformed rating represents the number of standard deviations which the raw rating is from that participant's mean rating.

- Z-transformation controls for potential individual scale bias (e.g. using only one or both extremes on a scale).

- Z-transformation makes direct comparison between studies that may not use the same scale easier (e.g. comparing responses on a binary scale [OK/not OK] vs. responses on a 7-point Likert scale)
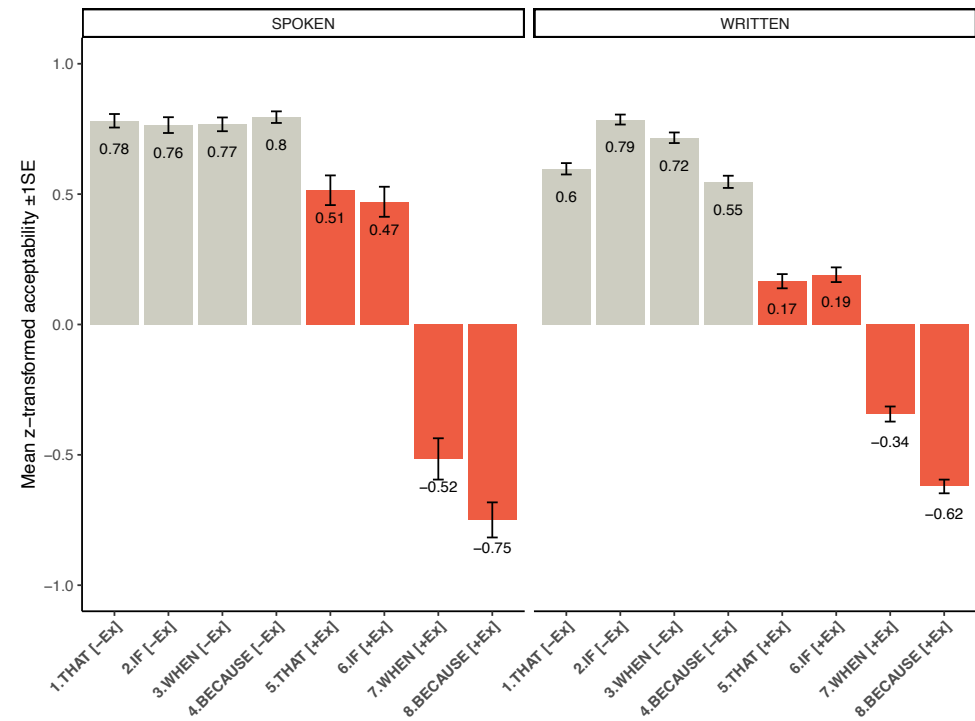
Modality

- **Interaction between TYPE and MODALITY?**

- **Only significant for Type 7**
  - *This is the exercise that I was surprised [when she actually completed __].*

- **Rated higher in the written (!) modality**

(marginal, p = 0.08)
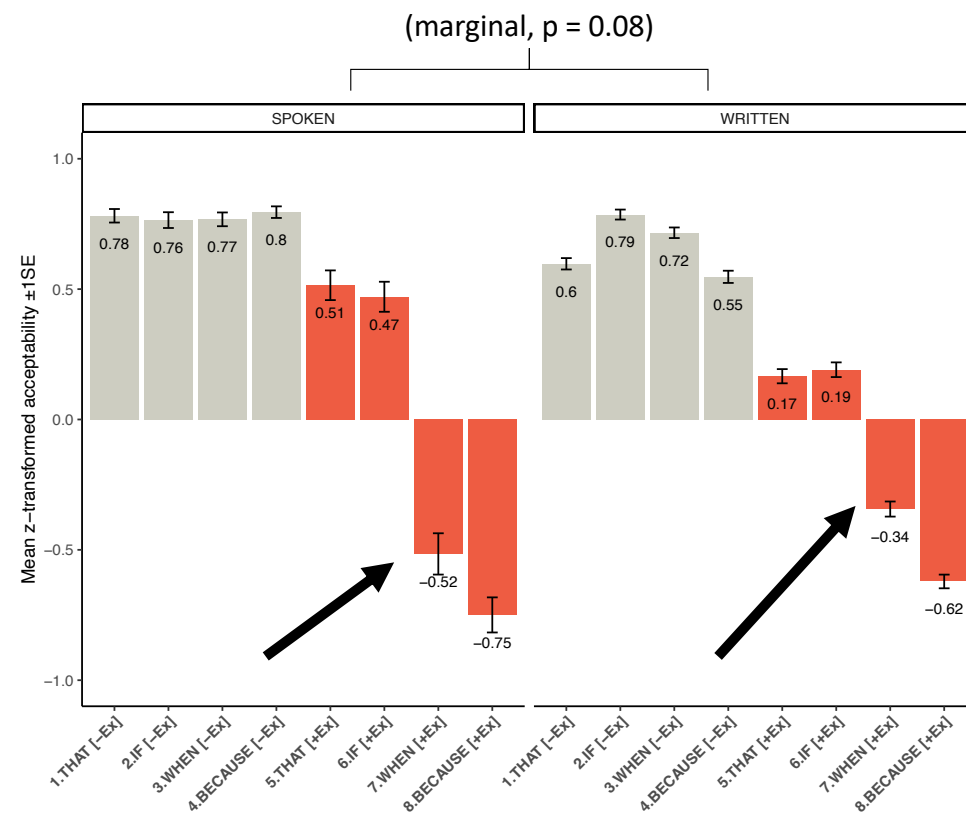
MODEL = lmer(ACC_z ~ TYPE*MODALITY + (1+TYPE+MODALITY|PARTICIPANT) + (1+MODALITY|ITEM), data=DATA_COMBINED)

(Participant) gender and (speaker) voice

- **No significant interactions with gender or voice (p > 0.35).**

- **No significant fixed 'main' effects voice (p > 0.8) or gender (p > 0.9).**

- **EXCEPT**:
  Speakers found extraction from *if-clauses* more acceptable when spoken with the female voice than with the male voice (p < 0.045).
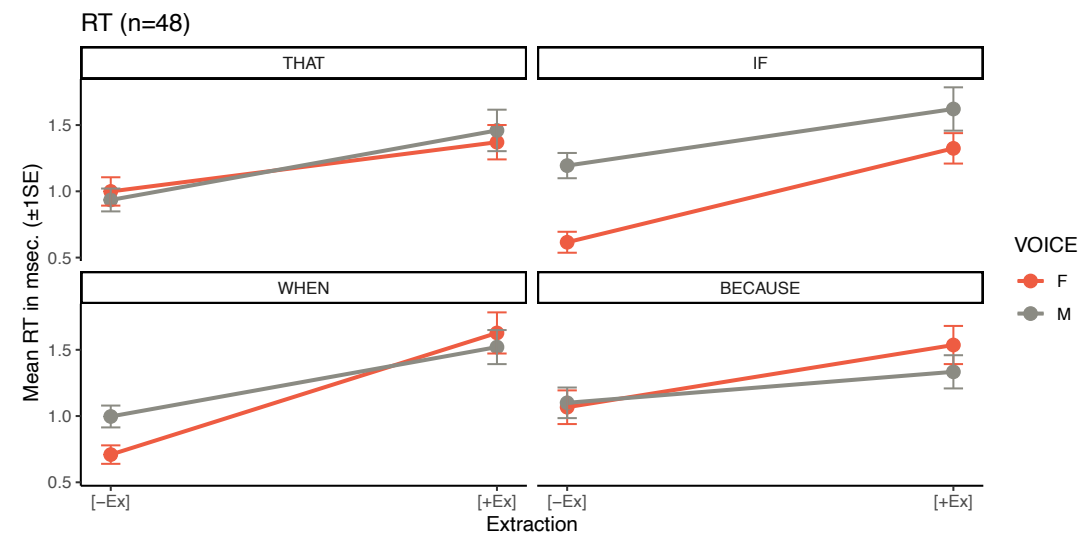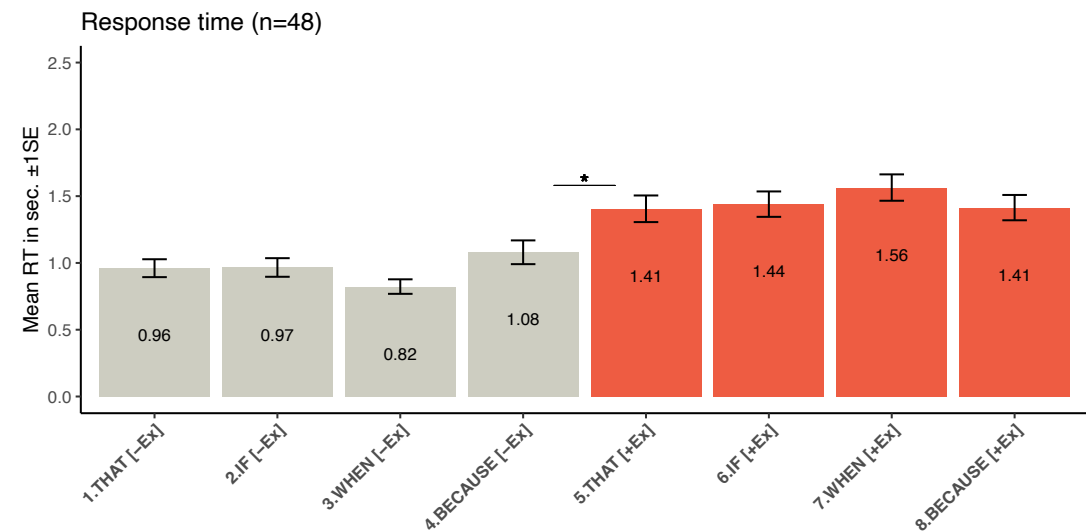 - *This is the exercise that I would be surprised [if she actually completed __].*

Acceptability (n=48)
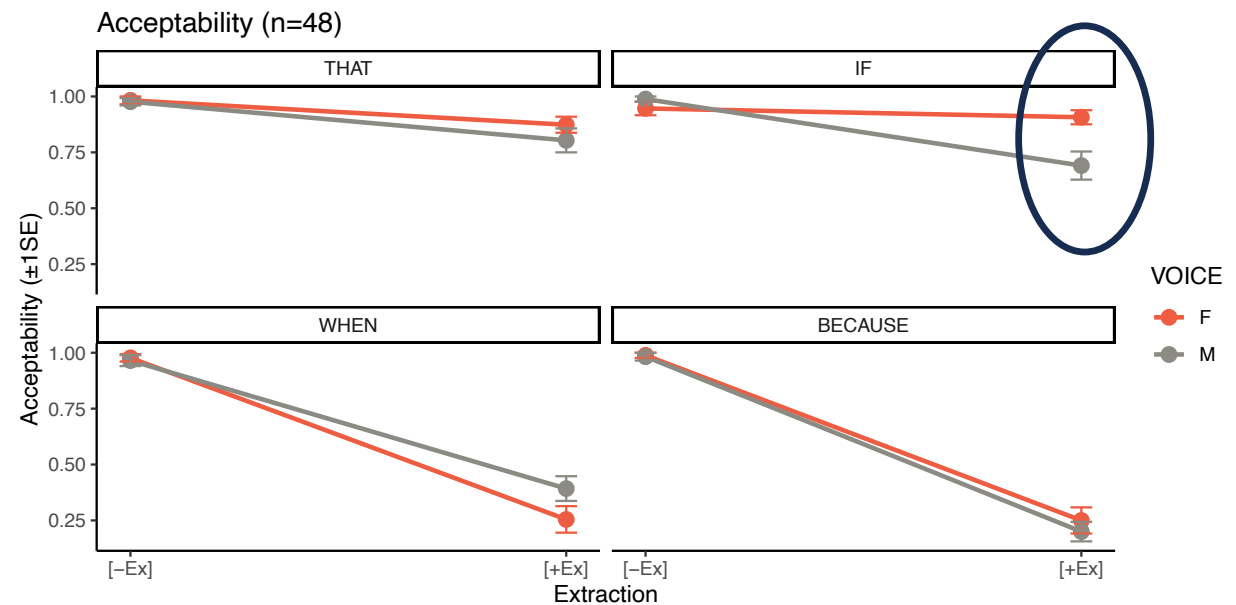
MODEL = glmer(ANSWER ~ VOICE*GENDER*TYPE + (1+VOICE+GENDER+TYPE|PARTICIPANT) + (1+VOICE+GENDER|ITEM), data = DATA1, family = "binomial")

Response time (RT)

- **Significant increase in RT with [+Ex] (p = 0.036)**
  - (on average 0.5 sec.)

- No significant interactions with gender or voice (p > 0.14).

- No significant fixed 'main' effects of voice (p > 0.91) or gender (p > 0.38).

Response time (n=48)

Mean RT in sec. ±1SE

1.THAT [−Ex]	2.IF [−Ex]	3.WHEN [−Ex]	4.BECAUSE [−Ex]	5.THAT [+Ex]	6.IF [+Ex]	7.WHEN [+Ex]	8.BECAUSE [+Ex]
0.96	0.97	0.82	1.08	1.41	1.44	1.56	1.41

RT (n=48)

Mean RT in msec. (±1SE)

VOICE
F
M

Conclusion

- <mark>Overall, the results are basically a replication of Nyvad et al. (2022), only in the spoken modality.</mark>

- Variation across the different adjunct clause types suggest that the CED does not hold in English and is not a universal constraint banning all extraction (cf. Nyvad et al. 2022).

- This holds for written as well as spoken sentences.
    - No modality difference – except of extraction from *when*, which is more acceptable in writing.
    - But all thigs being equal, such extraction is expected to be more acceptable in the spoken modality

- And across participant gender
    - No effect of participant gender or speaker voice,
    - Except for extraction from *if*, which was rated more acceptable with the female voice.

Nyvad, Anne Mette, Christiane Müller, and Ken Ramshøj Christensen. 2022. "Too True to Be Good? The Non-Uniformity of Extraction from Adjunct Clauses in English." *Languages* 7 (4): 244. https://doi.org/10.3390/languages7040244.

However

- Our study only involved two voices (one male and one female).

- Effects might be due to accidental properties of (one of the) the voices or of the AI-generated sequences.

- We must be very careful not to generalize to the population from the single significant voice effect in our study.

- Further studies with more voices are required to address this issue.
  - Work in progress… to be continued.

Thanks