Focus types and the prosody-gesture link in Catalan and German: A production study

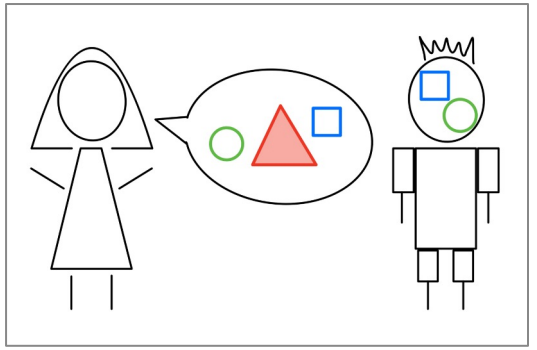
Alina Gregori, Paula G. Sánchez-Ramón, Pilar Prieto, Frank Kügler

IPrA conference Panel “Multimodal and prosodic markers of information structure and discourse structure” - Brussels 10.07.2023

Background

Focus

- Cognitive domain of information structure which indicates the **presence of alternatives that are relevant** for the interpretation of linguistic expressions (Krifka, 2008:247)



- Focus types:

Increase of pragmatic prominence

Baumann et al., 2006; Zimmermann, 2008

Background: non focused constituents
Information Focus: most important information
Contrastive Focus: notion of contrast to another element
Corrective Focus: disagreement to a previous statement

Video extracted from the M3D-TED corpus (Rohrer, 2022)

The points at which the satellite is closest to our antenna and the points at which it's furthest away

Video extracted from the M3D-TED corpus (Rohrer, 2022)

The points at which the satellite is CLOSEST to our antenna and the points at which it's furthest AWAY

Multimodal marking of focus

Video extracted from the M3D-TED corpus (Rohrer, 2022)

The points at which the satellite is CLOSEST to our antenna and the points at which it's furthest AWAY

Linguistic prominence:

Intonation

Body movements

Gesture and speech temporal integration

Prosody and gesture are **generally aligned in many languages** (semantically, pragmatically, phonologically) (McNeill, 1992)

- In **alignment**: the gesture **apex occurs at the peak of the rising pitch accent** (Esteve-Gibert & Prieto, 2013)
- In **prominence** patterns: **more prominent referents** were suggested to attract **more gestures and more prominent gestural movements** (e.g., Im & Baumann, 2020 for English; Debreslioska & Gullberg, 2022 for German)

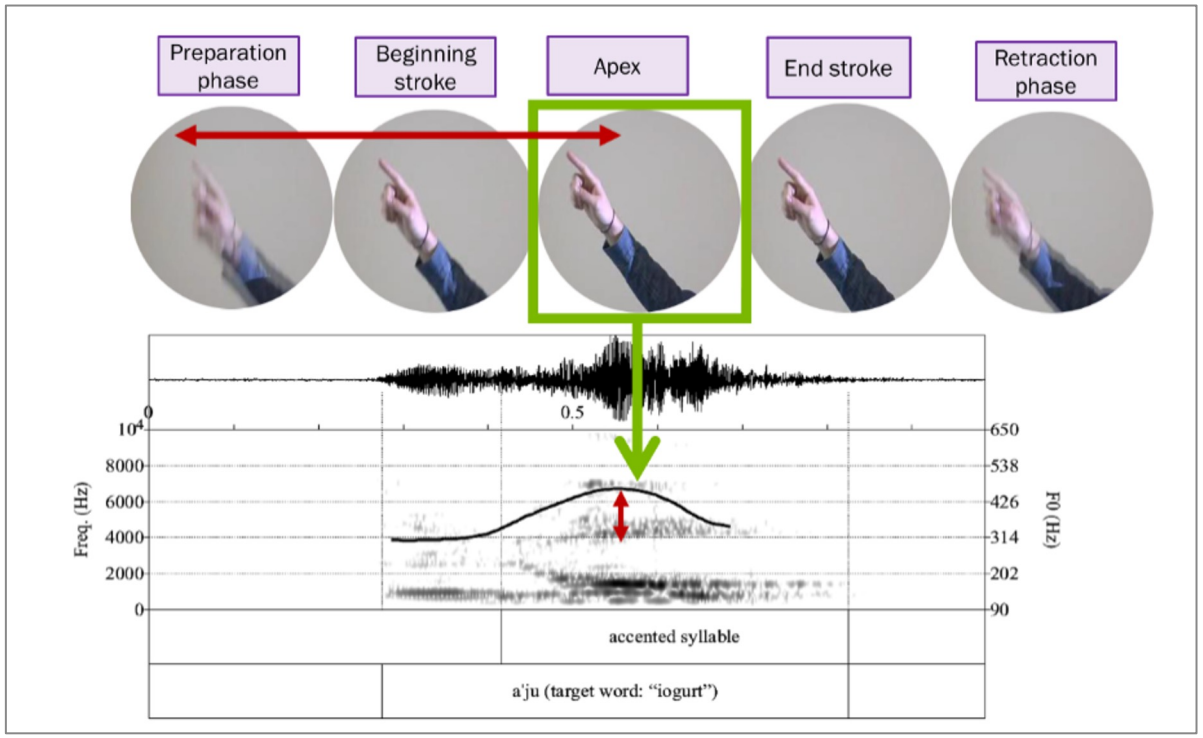
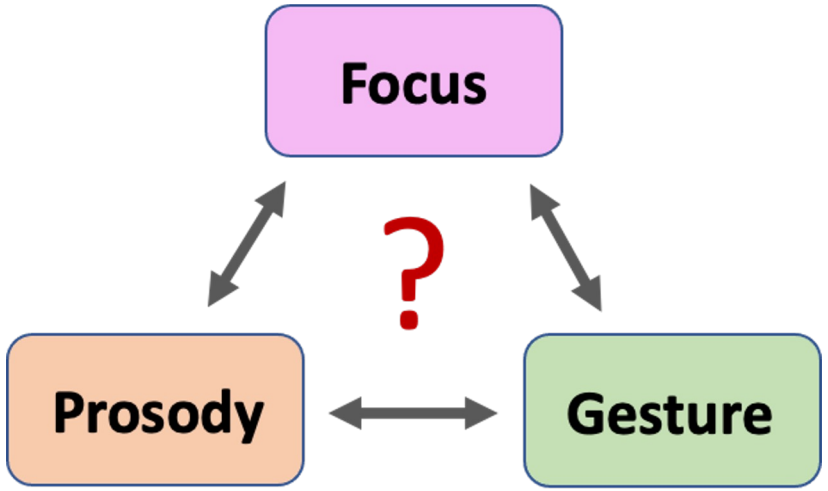


Figure adapted from Esteve-Gibert and Prieto (2013)

Interactions between focus, prosody and gesture

- Focused constituents receive **nuclear accentuation** (Féry & Kügler, 2008)
- Focused information that marks **contrast or correction** was suggested to carry **stronger levels of prosodic prominence than information focus conditions** (Zimmermann, 2008)
- **Non-referential (beat) gestures** (e.g., head nods, eyebrow movements, hand movements) have been reported to be **involved in discourse-marking functions** (e.g., Loehr, 2012)
- However, less is known about the **joint contribution of prosody and co-speech gestures to the marking of focus** conditions.



Research Questions

? Q1: **Is pragmatic prominence in focus types reflected in multimodal prominence?**

? Q2: **Is the relation between gesture use and focus types direct, or is it mediated by prosody?**

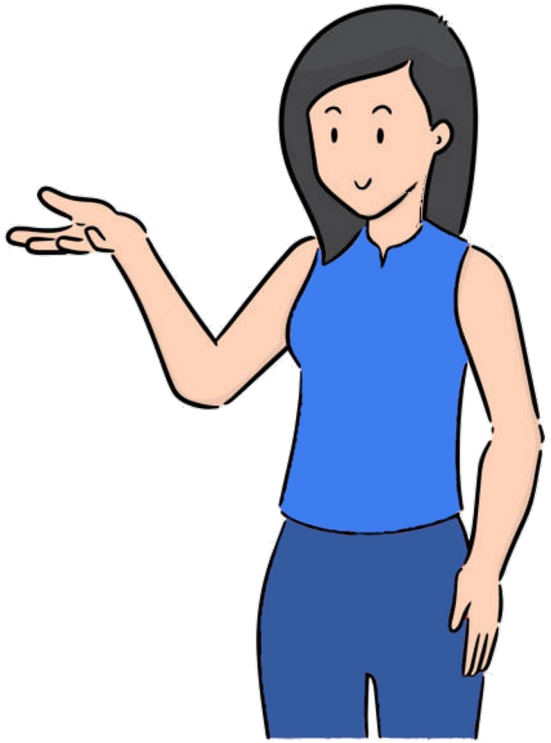
Why Catalan and German?

	German	**Catalan**
Prosodic typology	Intonation language	Intonation language
Rhythm class	Stress timed	Syllable timed
Accentuation patterns in the NP (Krahmer & Swerts, 2007)	Deaccentuation after nuclear pitch accent	Accentuation of all words within the NP
Prominence marking	→ Different language particular ways of prominence marking	

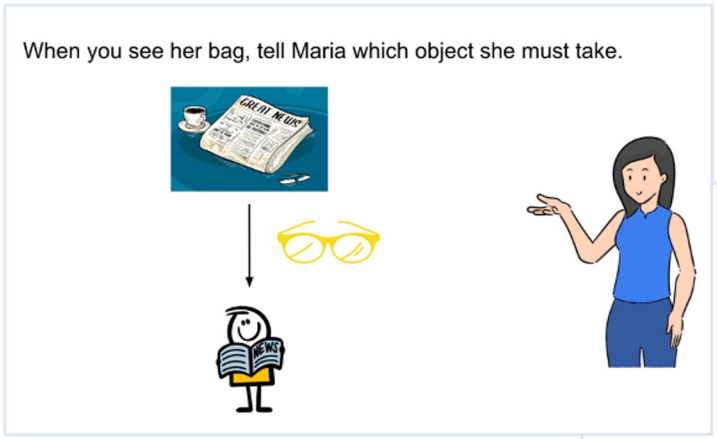
Our study

Methodology

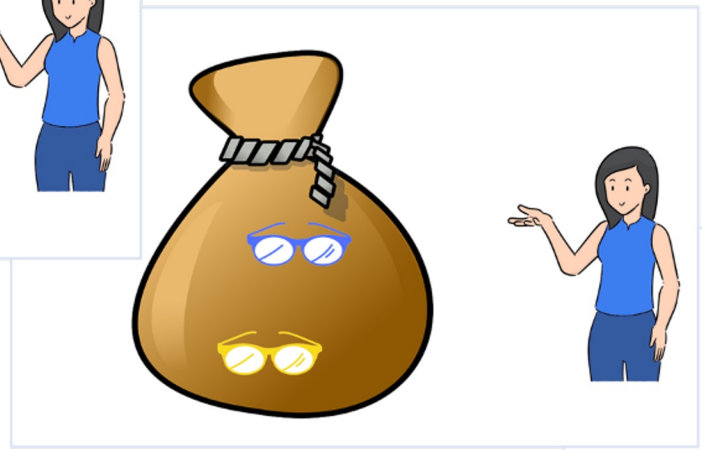
- **Focus elicitation task** adapted from Esteve-Gibert et al. (2021)
- Participants sit on a **high chair, in front of a screen**
- **They talk to a language learner, Maria,** in an interactive way
- She is **learning** the **colors** and **instructions**
- Their task is to **instruct her to take certain objects from a bag**



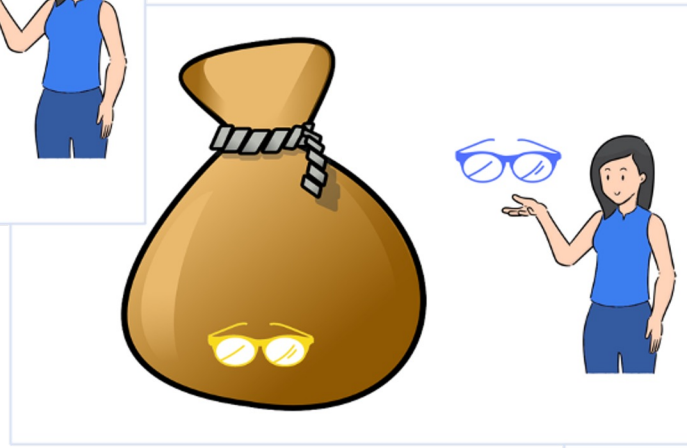
Methodology: example of contrastive and corrective conditions.



1. Context slide with target object



2. Object(s) prompted and instruction from the participant (contrastive condition).
'Maria, take the YELLOW glasses.'



3. Maria's response: she makes a mistake, and the participant corrects her (corrective condition).
'No, Maria, take the YELLOW glasses!'

4. Maria's response: she takes the correct object. **'Well done, Maria!'**

Methodology

- **14 participants** (7 German, 7 Catalan)
- Total of **392 Noun Phrases** (object + color of the object)
- Target domain for **focus is the adjective** (color)
- **Intended** items:
  - German: 7 speakers X 4 conditions X 7 items = 196 items
  - Catalan: 7 speakers X 4 conditions X 7 items = 196 items
- **Produced** items
  - **German: 187**
  - **Catalan: 181**

Data coding and analysis

- **Pitch accent presence** in Praat (Boersma & Weenink, 2022) following CatToBI and GToBI (Prieto et al., 2015; Grice et al., 2005)
- **Eyebrow movements, head movements** and **hand gestures (strokes)** in ELAN according to M3D (Rohrer et al., 2023)
- **Perceived prosodic prominence** from 0 to 3 (DIMA, Kügler et al., 2015, 2019, 2022)
- **Perceived gestural prominence:** from 0 to 3, as the **degree of visual saliency of the gesture** in relation to neighboring gestural movements, taking into account kinematic cues used in gesture production (e.g., size, speed, or "beat-like-ness" (Rohrer et al., 2023))
- **Analysis**:
  - **Number of cues** per focus condition
  - **Perceived prosodic and gestural prominence ratings** per focus condition
  - **Perceived general prominence** per focus condition (combination of prosodic and gestural prominence per item)

Examples - prosodic prominence ratings

0 1 2 3

Ger

*die gelbe [**Glühbirne**]^{T}*
*the yellow [**bulb**]^{T}*

*den [**gelben**]^{T} Farbtopf*
*the [**yellow**]^{T} paint*

*eine [**lilane**]^{T} Gabel*
*a [**purple**]^{T} fork*

*den [**lilanen**]^{T} Wanderrucksack*
*the [**purple**]^{T} backpack*

Cat

*el micròfon de color [**rosa**]^{T}*
*the [**pink**]^{T} microphone*

*el mòbil de color [**rosa**]^{T}*
*the [**pink**]^{T} mobilephone*

*el gerro de color [**rosa**]^{T}*
*the [**pink**]^{T} vase*

Examples - gestural prominence ratings

0 | 1 | 2 | 3

Results

Results – Number of cues (in percentages) produced by the participants on the adjectives (focused item) per condition, separated by language

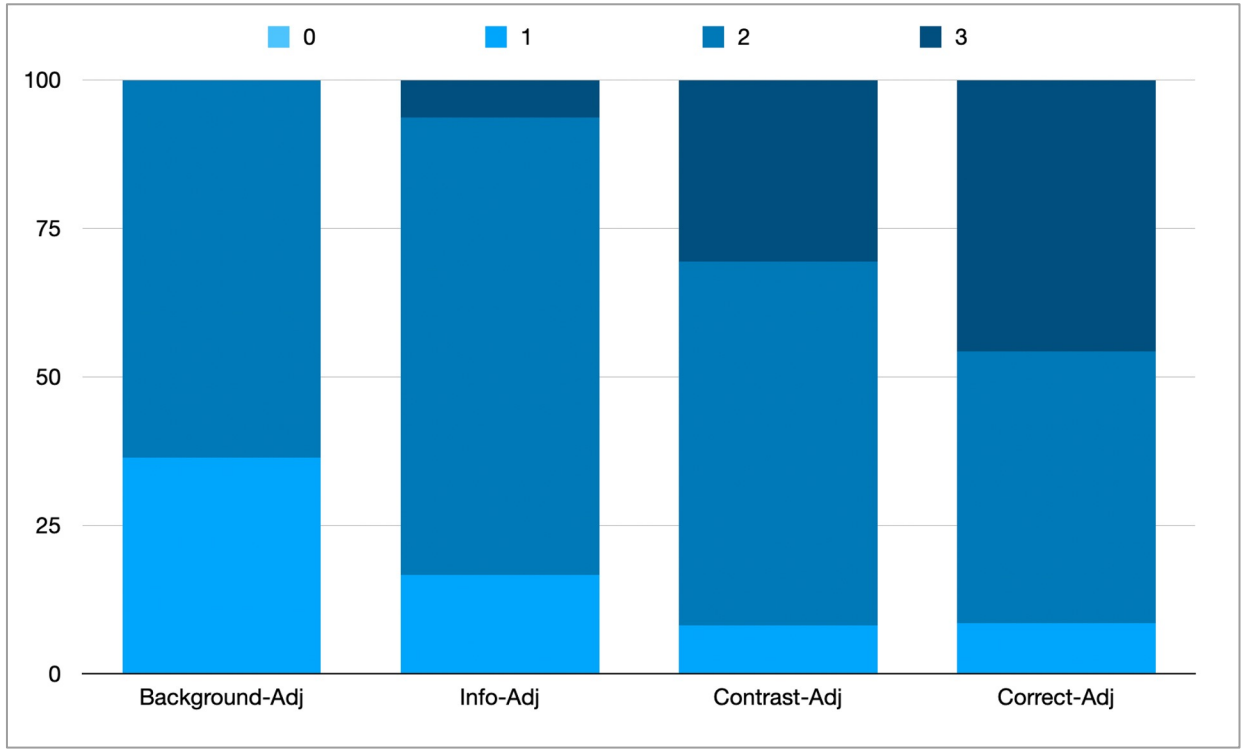
CUES		PITCH ACCENTS
GERMAN	BACK	97,7 %
	INFO	97,9 %
	CONT	97,9 %
	CORR	100 %
CATALAN	BACK	100 %
	INFO	100 %
	CONT	100 %
	CORR	100 %

Results – Perceived prosodic prominence

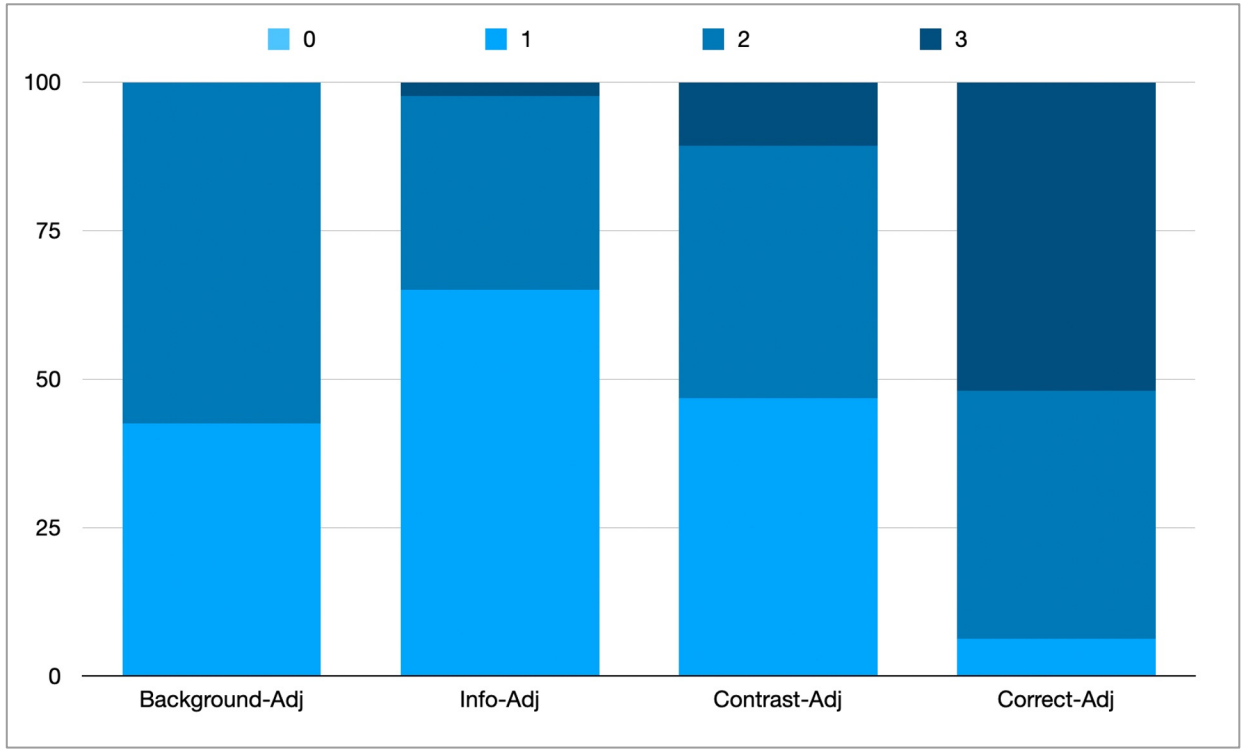
Increases with pragmatic prominence in both languages

- German: Increase from Background to Corrective, Contrastive and Corrective similar
- Catalan: Increase from Info to Corrective, stronger difference in Corrective, Background not in line

German



Catalan



Results – Number of cues (in percentages) produced by the participants on the adjectives (focused item) per condition, separated by language.

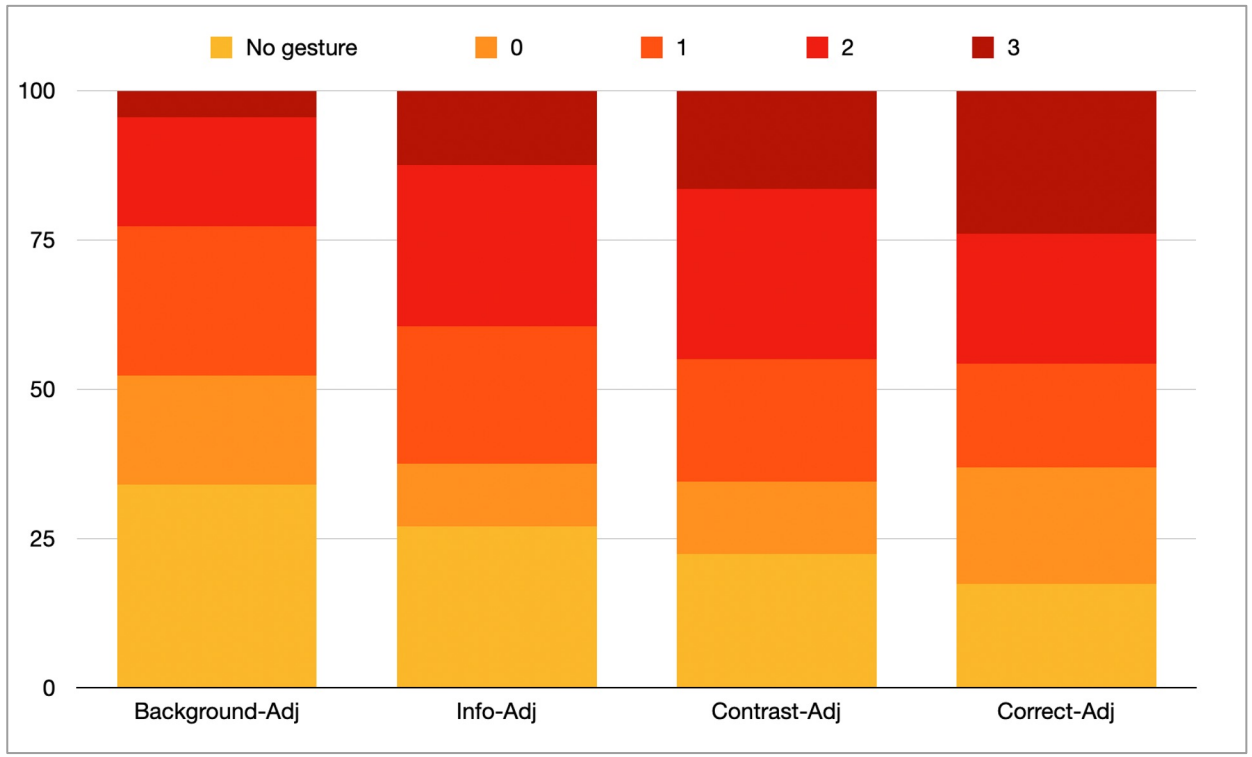
CUES		PITCH ACCENTS	MANUAL STROKES	HEAD MOVEMENTS	EYEBROW MOVEMENTS
GERMAN	BACK	97,7 %	34 %	38,6 %	13,6 %
	INFO	97,9 %	52 %	58,3 %	0 %
	CONT	97,9 %	55 %	67,3 %	14,3 %
	CORR	100 %	63 %	65,2 %	4,3 %
CATALAN	BACK	100 %	50 %	10 %	7,5 %
	INFO	100 %	39,1 %	28,2 %	6,5 %
	CONT	100 %	36,1 %	25,5 %	4,25 %
	CORR	100 %	66,6 %	41,6 %	20,8 %

Results – Perceived gestural prominence

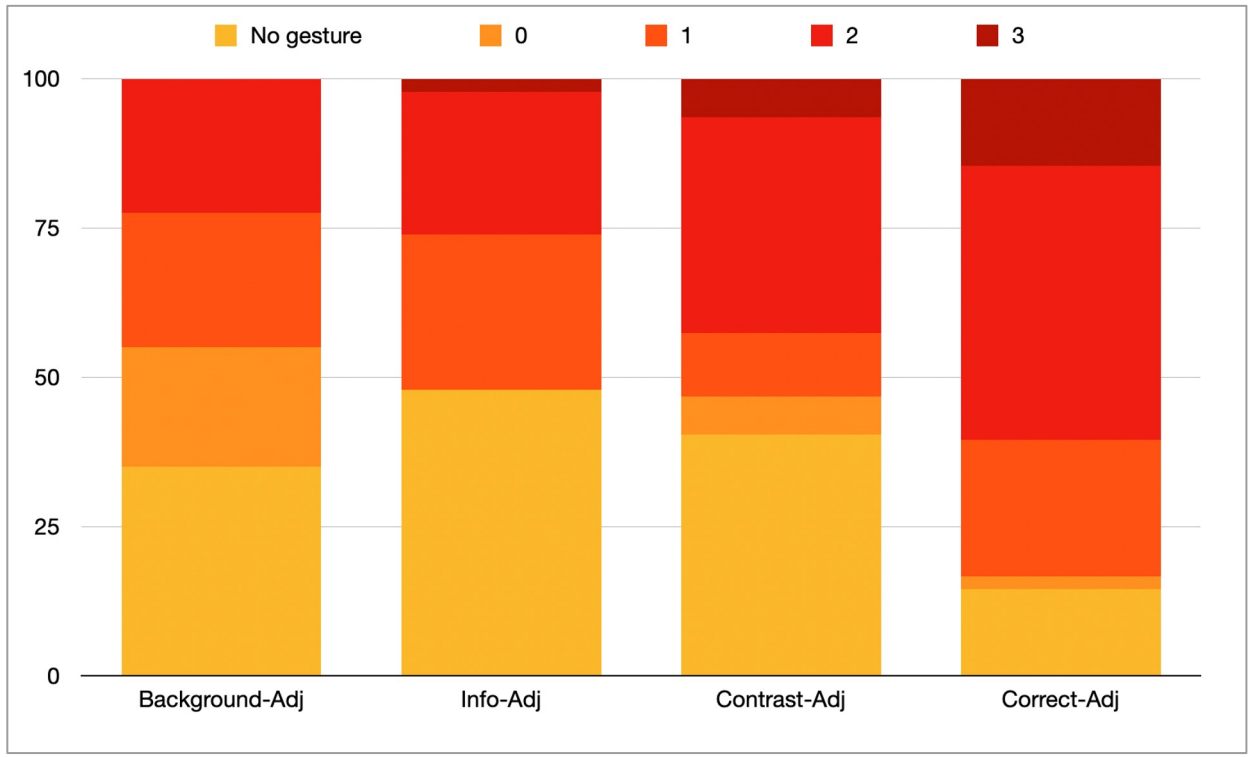
Increases with pragmatic prominence in both languages

- German: stepwise increase, not too big differences
- Catalan: stepwise increase, stronger prominence in Corrective

German



Catalan

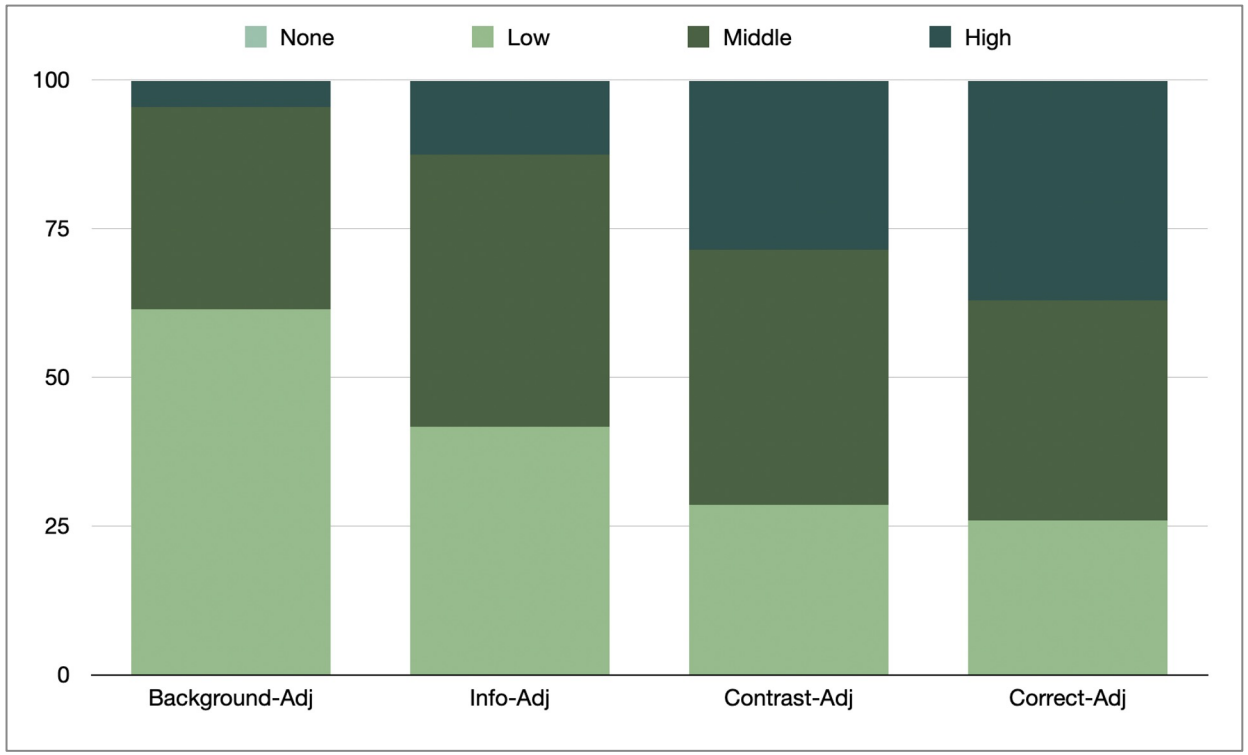


Results – Perceived general prominence (combination of prosodic and gestural prominence in each item)

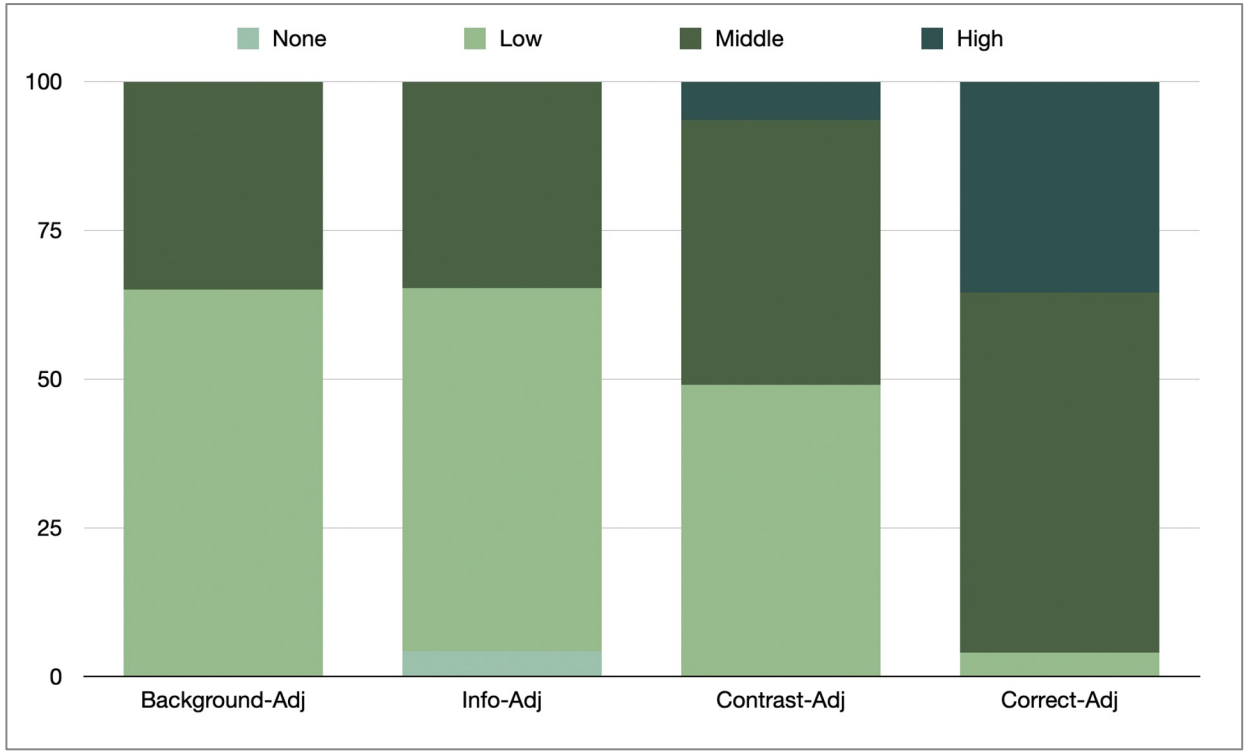
Increases with pragmatic prominence in both languages

- German: stepwise increase, but Contrastive and Corrective are quite similar
- Catalan: similarity between Background, Info and Contrastive, while Corrective stronger
- More prominent categories tend to be marked by both acoustic and gestural prominence

German



Catalan



Discussion

+ Conclusion

Conclusions - preliminary

? Q1: **Is pragmatic prominence in focus types reflected in multimodal prominence in Catalan and German?**

- For **prosody**: it is not reflected when we look at pitch accent presence, but **it is when we look at prosodic perceived prominence**.
- For **gesture**: we can see a tendency for the number of gestures (manual and head movements) to increase while pragmatic prominence increases, but it gets **clearer when we look at gestural perceived prominence.**

? Q2: **Is the relation between gesture use and focus types direct, or is it mediated by prosody (in Catalan and German)?**

- From preliminary data, we can see that **gestural cues tend to mark focus categories**, as prosodic cues do.

Discussion

- The method elicits **natural gestures while enabling to control for focus**
- **Number of gestures** per condition shows an i**ncreasing tendency** with more prominent focus categories
- **Perceived prominence increases** both **visually and acoustically on the most pragmatically prominent constituents** (information < contrastive < corrective)
- Differences in the prominence gradations on focus types between Catalan and German
- Analyse **gestural prominence cues**: amplitude, velocity, size…
- Analyse **acoustic prominence cues**: intensity and duration of the accented syllables, pitch accent height
- Analyse **temporal alignment** of hand gesture apexes and pitch accents
- Next steps:
  - Finish data collection (40 participants/language)
  - Gesture factors: referentiality, number of gestural cues at the same time
  - Insight about the main prominence patterns in the whole TP (noun + adjective)

Our team: thanks!

Alina Gregori

Prof. Dr. Frank Kügler

Prof. Dr. Pilar Prieto

Paula G. Sánchez-Ramón

Thanks to our Student assistants Leoni Hahn, Jonas Pfeffer, and Carla Rufí for help in data annotation!

DFG Deutsche Forschungsgemeinschaft
ViCom
GOETHE UNIVERSITÄT FRANKFURT AM MAIN
upf. Universitat Pompeu Fabra Barcelona
AGAUR Agència de Gestió d'Ajuts Universitaris i de Recerca
GOBIERNO DE ESPAÑA MINISTERIO DE ECONOMÍA Y COMPETITIVIDAD

Thank you very much!

References

Baumann, Stefan, Martine Grice & Susanne Steindamm. 2006. Prosodic Marking of Focus Domains - Categorical or Gradient? 301–304.

Boersma, Paul & David Weenink. 2022. Praat: doing phonetics by computer [Computer program]: Version 6.2.09 (http://www.praat.org/).

Debreslioska, Sandra & Marianne Gullberg. 2022. Information Status Predicts the Incidence of Gesture in Discourse: An Experimental Study. Discourse Processes 59(10). 791-827.

Esteve-Gibert, Núria & Pilar Prieto. 2013. Prosodic structure shapes the temporal synchronization of intonation and manual gesture movements.

Esteve-Gibert, N., Lœvenbruck, H., Dohen M. & D'Imperio, M. (2021). Pre-schoolers use head gestures ratherthan prosodic cues to highlight important information in speech. Developmental Science. e13154.

Féry, Caroline & Frank Kügler. 2008. Pitch accent scaling on given, new and focused constituents in German. Journal of Phonetics 36(4). 680–703.

Grice, Martine, Stefan Baumann & Ralf Benzmüller. 2005. German Intonation in Autosegmental-Metrical Phonology. In Sun-Ah Jun (ed.), Prosodic Typology: The Phonology of Intonation and Phrasing, 55–83. Oxford: Oxford University Press.

Krahmer, Emiel & Mark Swerts. 2007. The effects of visual beats on prosodic prominence: Acoustic analyses, auditory perception and visual perception. Journal of Memory and Language 57. 396-414. Elsevier.

Krifka, Manfred. 2008. Basic notions of information structure. Acta Linguistica Hungarica 55(3). 243–276.

Kügler, Frank, Bernadett Smolibocki, Denis Arnold, Stefan Baumann, Bettina Braun, Martine Grice, Stefanie Jannedy, Jan Michalsky, Oliver Niebuhr, Jörg Peters, Simon Ritter, Christine T. Röhr, Antje Schweitzer, Katrin Schweitzer & Petra Wagner. 2015. DIMA - Annotation guidelines for German intonation. 317.

Kügler, Frank, Stefan Baumann, Bistra Andreeva, Bettina Braun, Martine Grice, Jana Neitsch, Oliver Niebuhr, Jörg Peters, Christine T. Röhr, Antje Schweitzer & Petra Wagner. 2019. Annotation of German Intonation: DIMA compared with other annotation systems. 1297-1301.

Kügler, Frank, Stefan Baumann & Christine T. Röhr. 2022. Deutsche Intonation, Modellierung und Annotation (DIMA): Richtlinien zur prosodischen Annotation des Deutschen. In Cordula Schwarze & Sven Grawunder (eds.), Transkription und Annotation gesprochener Sprache und multimodalerInteraktion: Konzepte, Probleme, Lösungen, 23–54. Tübingen: Narr Francke Attempto Verlag.

Im, Suyeon & Stefan Baumann. 2020. Probabilistic relation between co-speech gestures, pitch accents and information status. Proceedings of the Linguistic Society of America 5(1). 685697.

Loehr, Daniel P. 2012. Temporal, structural, and pragmatic synchrony between intonation and gesture. Laboratory Phonology: Journal of the Association for Laboratory Phonology 3(1). 71–89.

McNeill, David. 1992. Hand and mind: What gestures reveal about thought. Chicago: University of Chicago Press.

Prieto, P., Borràs-Comes, J., Cabré, T., Crespo-Sendra, V., Mascaró, I., Roseano, P., Sichel-Bazin, R., & Vanrell, M.M. (2015). Intonational phonology of Catalan and its dialectal varieties. In S. Frota & P. Prieto (Eds.), Intonation in Romance, 9-62. Oxford: OUP.

Rohrer, Patrick. 2022. A temporal and pragmatic analysis of gesture-speech association. PhD thesis. UPF, Barcelona, Spain.

Rohrer, Patrick L., Ingrid Vilà-Giménez, Júlia Florit-Pons, G. Gurrado, Núria Esteve-Gibert, Ada Ren, Stefanie Shattuck-Hufnagel & Pilar Prieto. 2023. The MultiModal MultiDimensional (M3D) labeling system for the annotation of audiovisual corpora: Gesture Labeling Manual. UPF Barcelona.

Zimmermann, M. (2008). Contrastive focus and emphasis. Acta Linguistica Hungarica, 55, 347–360.

Linguistic prominence

> Prominence as a relational property. […] Prominence entails the property of being an ‘organizational principle’ for linguistic structure or, in other words, that a prominent element organizes its environment, providing a structure for the context in which it appears […].
>
> Grice & Kügler, 2021:253

Acoustic Prominence Marking

Prosodic Prominence

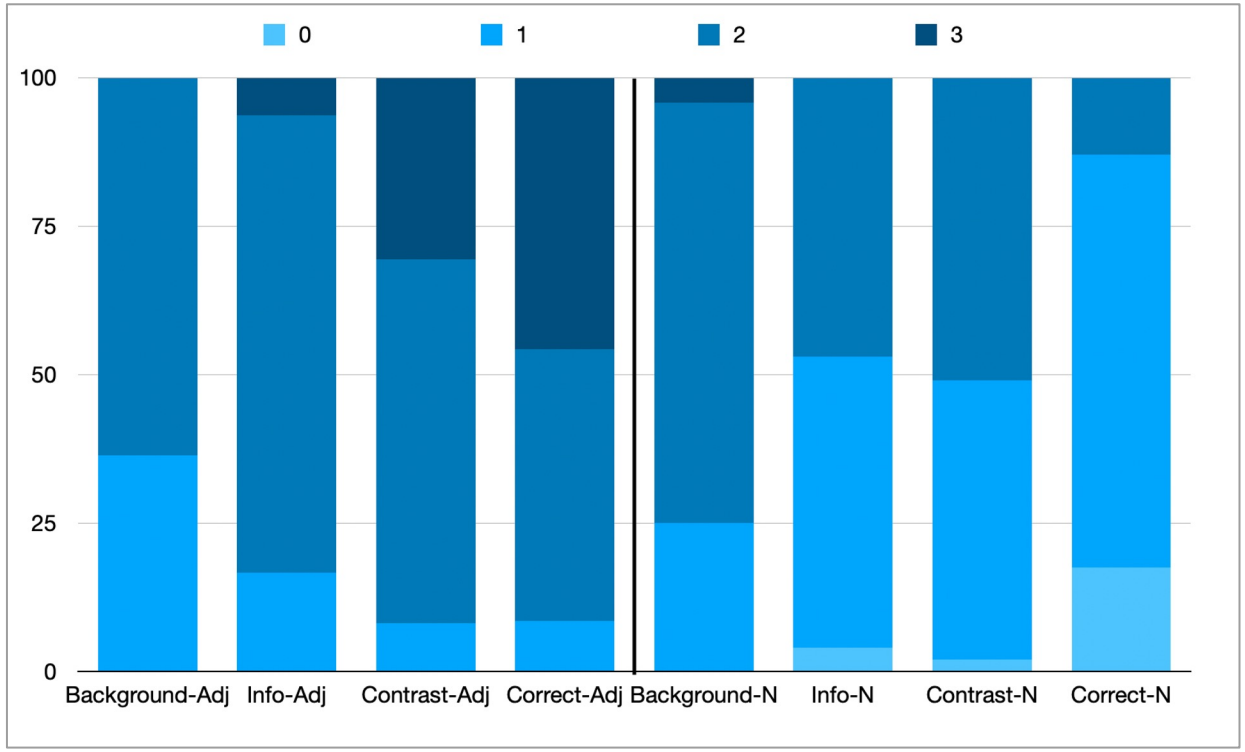
- Pitch accents can be used as indicators for prominence Ladd, 2008
- Higher f0 → higher prominence Baumann & Röhr, 2015; Kügler & Calhoun, 2020
- Pitch accent prominence scale (GER; Baumann & Röhr, 2015)

Perceived Prominence

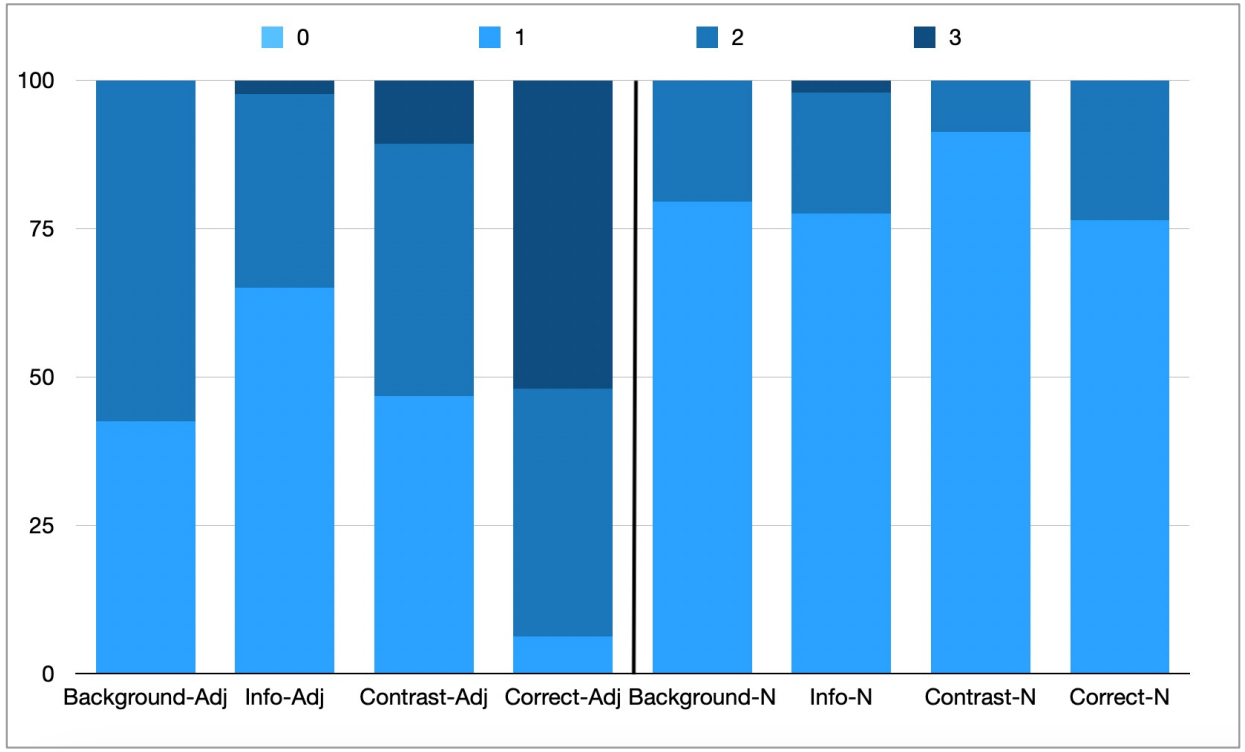
- Clearness of sounds and preciseness of pitch contours contribute to prominence perception Lindblom, 1990
- Measures like intensity, duration or contour of the acoustic signal contribute to prominence perception
- Do not mistake visual salience for perceptual reality, taking production AND perception into account Niebuhr & Reetz, 2020
- Strategies for reliable prominence ratings (e.g., Rapid Prosody Transcription; identification of individual cues to the contrastive prosodic elements of an utterance Cole & Shattuck-Hufnagel, 2016)

Results – Perceived prosodic prominence including nouns

German

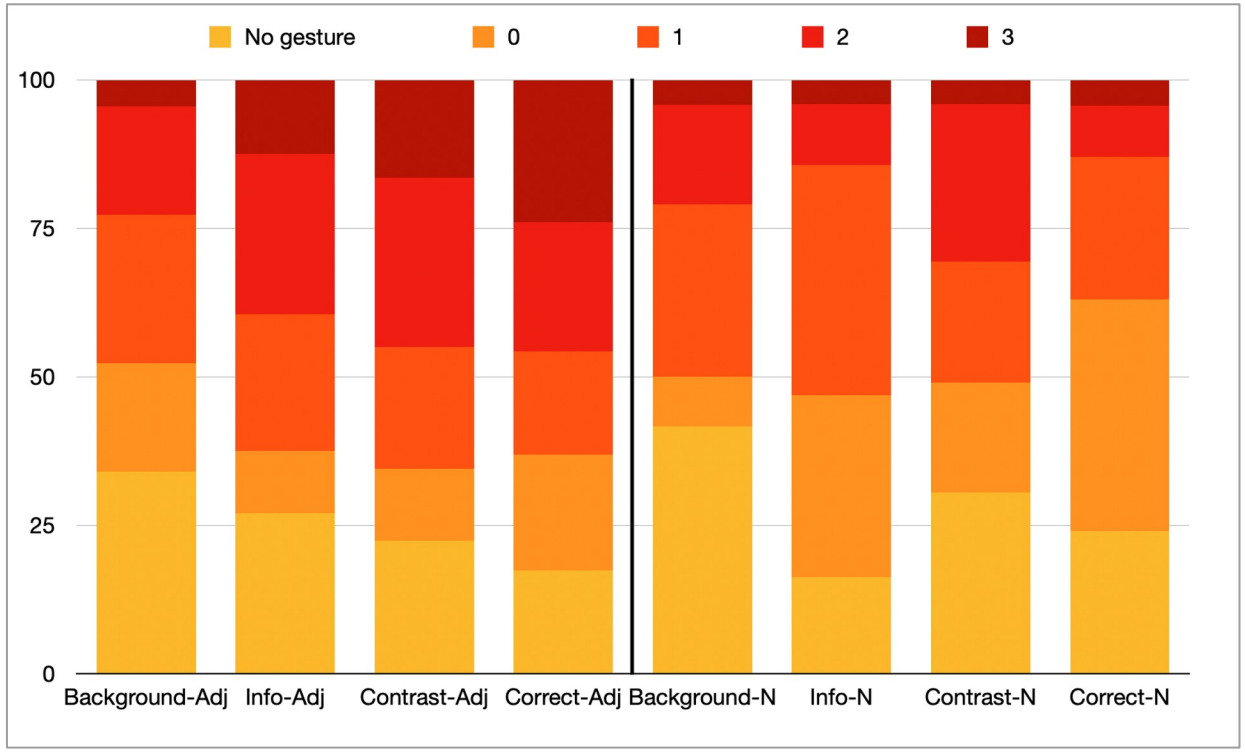


Catalan

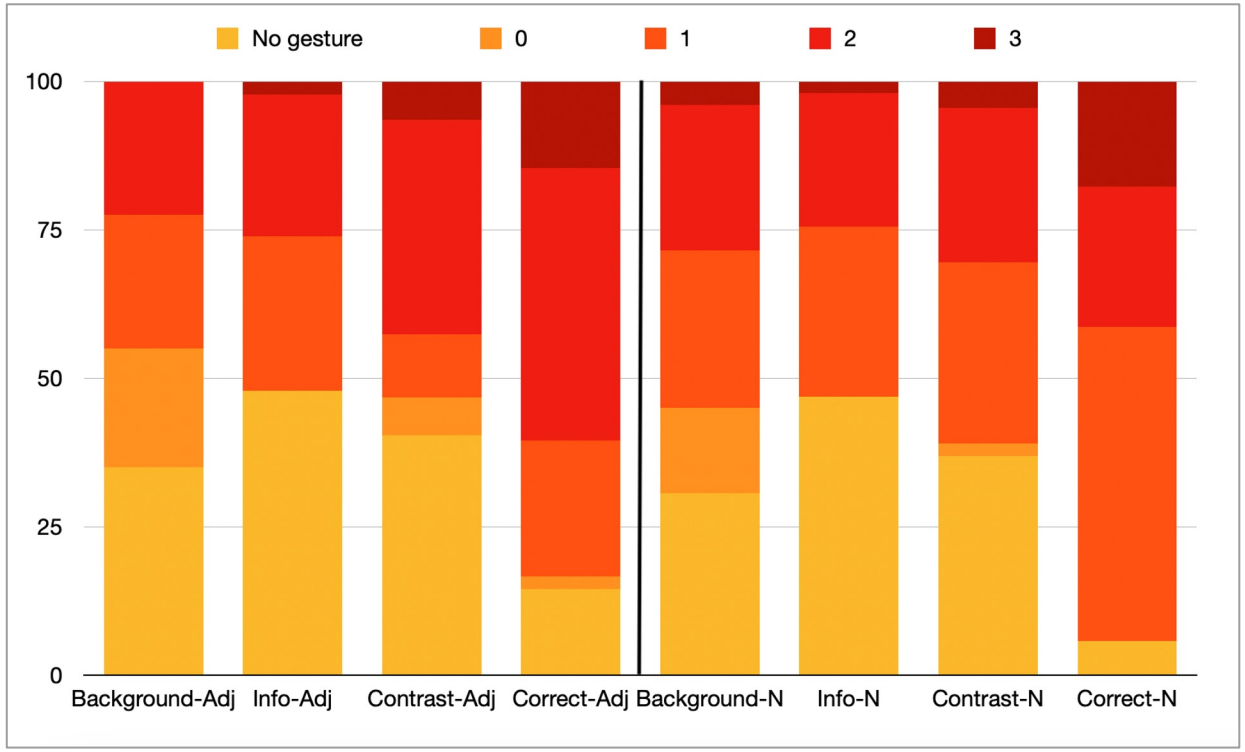


Results – Perceived visual prominence including nouns

German

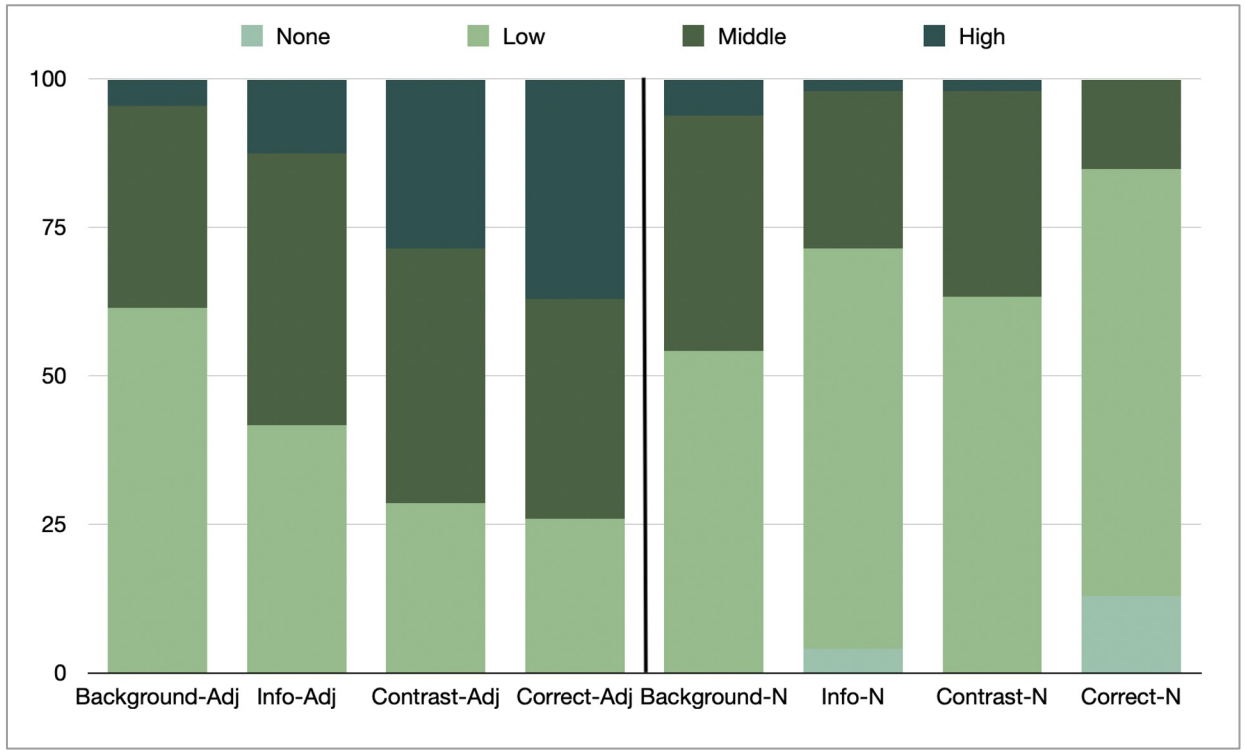


Catalan

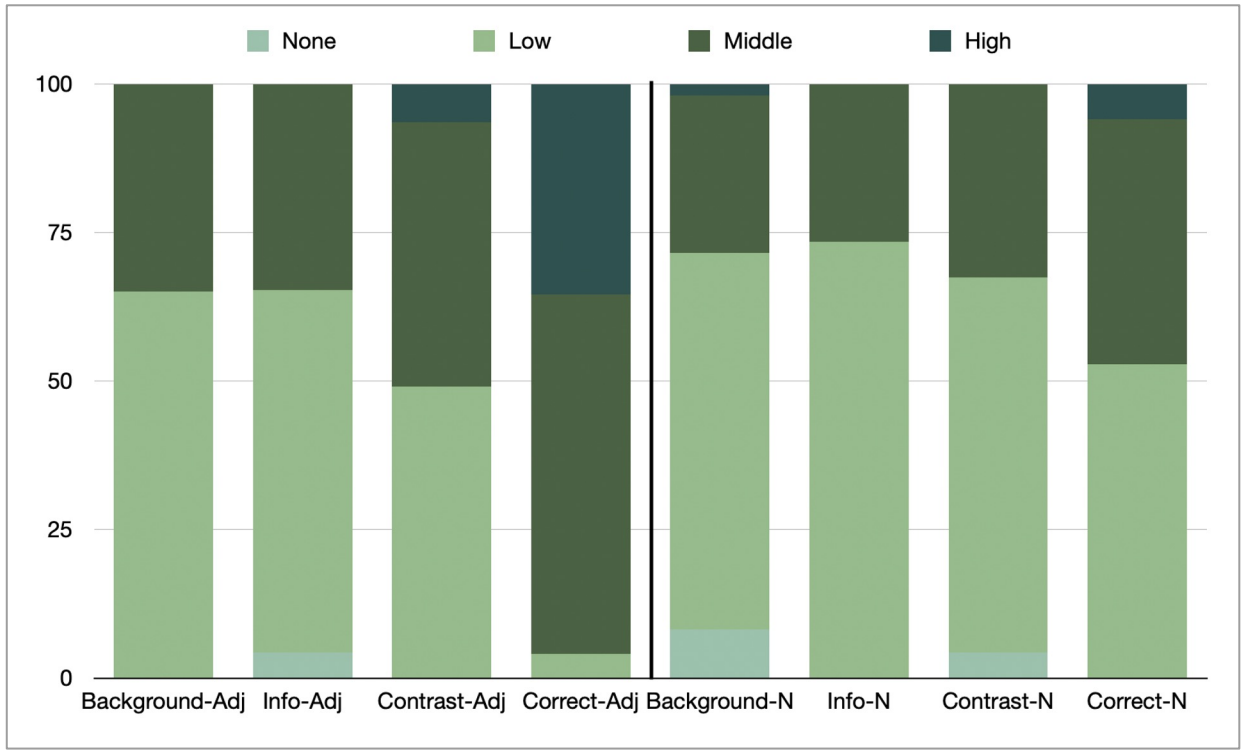


Results – Perceived combined prominence including nouns

German



Catalan



Items

German

*… aus dem Beutel die [**gelbe**]^{T} Brille nehmen*

Catalan

*Molt bé, Maria, has d'agafar les ulleres de color [**groc**]^{T}*

*Good, Maria, you need to take the [**yellow**]^{T} glasses*

Data Coding and Analysis