The MultiFocus corpus:
An audiovisual database for the elicitation of focus types and semi-spontaneous speech and gestures in Catalan and German

*Alina Gregori[1], *Paula G. Sánchez-Ramón[1,2], Pilar Prieto[3,2] and Frank Kügler[1]
* Shared first authorship
[1]Goethe University Frankfurt, [2]Universitat Pompeu Fabra, [3]ICREA

Full citation: Gregori, A., Sánchez-Ramón, P. G., Prieto, P., & Kügler, F. (2025). The MultiFocus corpus: An audiovisual database for the elicitation of focus types and semi-spontaneous speech and gestures in Catalan and German. https://doi.org/10.17605/OSF.IO/4M53Q

Short citation: Gregori, A., Sánchez-Ramón, P. G. et al. (2025). The MultiFocus corpus: An audiovisual database for the elicitation of focus types and semi-spontaneous speech and gestures in Catalan and German. https://doi.org/10.17605/OSF.IO/4M53Q

Last updated: May 2025
Created: March 2025

The MultiFocus corpus: A description

Table of contents

The MultiFocus corpus: A description

This document contains a general description of the full MultiFocus corpus. The MultiFocus corpus is an audiovisual database for the elicitation of focus types and semi-spontaneous speech and gestures in Catalan and German. More detailed information on materials and annotations for each language, namely Catalan and German, can be found in the OSF subcomponents (see Table 5 in section 7 for information about the organization of the files).

- MultiFocus corpus - CAT: https://osf.io/7z524/
- MultiFocus corpus - GER: https://osf.io/86zeg/

1. On ethical aspects of the corpus

Ethical approval for the data collection of this corpus was granted to the Linguistics Lab at Goethe University Frankfurt by the German Linguistics Society Ethics Commission under No. 2024-03. Before starting the task and the recording, participants consented to taking part in the data collection and to being video-recorded by signing a consent form.

2. Description and structure of the MultiFocus corpus

This audiovisual corpus was designed with the purpose of investigating the **multi**modal marking of **focus** (MultiFocus) **types** (information focus, contrastive focus, and corrective focus, based on Krifka, 2008) with regard to prosody and gesture, in Catalan and German.

Focus is a cognitive domain of information structure that represents choosing a referent from a set of alternatives (Krifka, 2008). We distinguish between three types of focus which increasingly convey pragmatic meaning (Repp, 2014), likely leading to diverse realizations of prosodic and gestural cues: **Information focus** marks most important information in the discourse, **contrastive focus** evokes the notion of contrast to another element and **corrective focus** shows disagreement to a previous statement.

The corpus is based on a focus elicitation task in which participants were asked to follow a visual display to instruct a language learner to take certain items from a bag, which are manipulated to elicit different focus types. The procedure was inspired by the focus elicitation task used in Esteve-Gibert et al. (2022) on children and adapted to adults so that we could obtain multimodal realizations of focus types. Participants were prompted to produce target phrases containing a noun and a color adjective with narrow focus on the adjective (see more details on the method in subsection 3.4 below). A summary of the target productions elicited in each focus type is displayed in Table 1. The recording sessions targeted 245 sentences per focus type per language (7 per participant). A total number of 702 focused items were produced in Catalan (220 information, 241 contrastive, 241 corrective), and a total of 682 in German (223 information, 236 contrastive, 223 corrective).

Table 1: *Overview of the elicited focus conditions in the MultiFocus corpus per language, including exemplary target productions and items visually displayed in the task. The focused word is marked in square brackets and the target phrase marked in bold.*

Language	Focus Condition	Target production	Items displayed
Catalan	Information	Maria, agafa la **forquilla [lila]**^{F}. *(Maria, take the [purple]^F fork.)*	purple fork
	Contrastive	Maria, agafa la **jaqueta [negra]**^{F}.	black jacket

		(Maria, take the [black]^F jacket.)	blue jacket
	Corrective	No, Maria, agafa la **jaqueta [negra]^F**. (No, Maria, take the [black]^F jacket.)	black jacket blue jacket
German	Information	Maria, nimm die **[lila]^F Gabel** aus dem Beutel. (Maria, take the [purple]^F fork from the bag.)	purple fork
	Contrastive	Maria, nimm die **[grüne]^F Jacke** aus dem Beutel. (Maria, take the [green]^F jacket from the bag.)	green jacket blue jacket
	Corrective	Nein, Maria, nimm die **[grüne]^F Jacke** aus dem Beutel. (No, Maria, take the [green]^F jacket from the bag.)	green jacket blue jacket

The MultiFocus corpus also includes two additional conditions with a similar structure as the experimental conditions. The first one displays two objects that only differ in their color and, after the participant's instruction, Maria takes the correct object. This condition was included so that Maria didn't always make a mistake after seeing two objects in her bag. It serves as a filler/distractor item, as it has a similar structure to contrastive focus but it is not followed by a mistake and thus, a correction by the participant. The second of these conditions had the goal to additionally introduce an action to be performed with an object. There are 490 productions of these conditions per language in the corpus (245 Contrastive fillers and 245 action items in Catalan; 245 Contrastive fillers and 245 action items in German), and examples for these productions are shown in Table 2.

Table 2: *Overview of the additionally recorded conditions in the MultiFocus corpus per language, including exemplary target productions and items visually displayed in the task. The focused word (in case of contrastive fillers) or the action (in case of the action item) is marked in square brackets and the target phrase marked in bold.*

Language	Condition	Target production	Items displayed
Catalan	Contrastive Filler	Maria, agafa el **gerro [lila]^F**. (Maria, take the [purple]^F vase.)	purple vase black vase
	Action items	Maria, agafa la **camisa groga** [per a fer una presentació]^A. (Maria, take the yellow shirt [to do a presentation]^A)	yellow shirt "presenting" action "baking cake" action
German	Contrastive Filler	Maria, nimm die **[lila]^F Vase** aus dem Beutel. (Maria, take the [purple]^F vase from the bag.)	purple vase green vase
	Action items	Maria, nimm die **gelbe Bluse** [zum Präsentieren]^A aus dem Beutel. (Maria, take the yellow shirt [for a presentation]^A from the bag.)	yellow shirt "presenting" action "baking cake" action

In the full recording session, each participant was prompted to produce 35 target productions which are included in the corpus (seven per sentence type), leading to 1225 sentences per language. In total, participants produced 1223 sentences in Catalan and 1224

sentences in German, which sums up to 2447 sentences contained in the corpus. Lists of all prompted productions per language can be found in the descriptions of the language subcomponents of the OSF. See the introduction of this document to find the links to the OSF subcomponents.

3. Data collection methods

3.1 Participants

The corpus consists of 70 speakers in total, 35 native speakers of Catalan (23 female, 12 male; mean age = 21,5 years) and 35 native speakers of German (20 female, 14 male, 1 non-binary; mean age = 27,8 years) respectively. The Catalan participants were recorded in Barcelona, 83% of them indicated themselves as being Catalan-dominant, and 74% of them reported a daily use of Catalan above 75% of the time. The German participants were all recorded in Frankfurt am Main and reported to be monolingual native speakers of German. All 70 participants were healthy adults, none of the speakers reported any speech or hearing disorders and were paid 10 € for their participation. All participants consented to being video-recorded for scientific purposes.

3.2 Setting

Recordings were conducted in rooms at Universitat Pompeu Fabra and Goethe University Frankfurt with the following recording set-up: A high chair was prepared for the participants in front of a digital screen in which they saw the stimuli. The high chair was chosen to create a relaxed body position without physical restrictions for participants to gesture. Participants were recorded using one camera (Sony HDRCX625 Camcorder) which was positioned with a slight diagonal angle and an external microphone (RODE Wireless Go for Catalan, Sennheiser MD46 for German). The experimenter sat in the same room behind the participants to not interfere in the communicative situation, and controlled the stimuli presentation from a laptop. Figure 1 shows the experimental setup in the lab in Barcelona and Frankfurt.

Figure 1: *Recording setup including screen, camera, and high chair in the lab in Barcelona (left) and Frankfurt (right).*

3.3 Materials: the stimuli presentation

The experiment's stimuli presentation introduced a digital character called Maria whom the participants had to interact with. She was located at the right side of the screen. A total of 35 target items were displayed in the stimuli presentation, see the full list of items for each

language in the respective subcomponents. For items in the conditions information focus, contrastive focus, corrective focus and contrastive fillers (item 1 to 28), an initial display (first picture in Figures 2 - 3) in the centre of the screen showed the target object that Maria had to take as well as a representation of the action that the object was used for, and a second display showed a bag containing the target object plus the possible competitor objects depending on the condition elicited. The objects were never presented as written text, but only as pictures.

For the action items (item 29 to 35), the same slide introducing the target object was displayed (first picture in Figure 4), but then instead of two competing objects, two competing actions were shown. The participants were asked to choose the action to be performed with the target object, while still mentioning the object and its color. For exemplary items, see Tables 1 and 2, and for illustrations of materials of the different conditions, see Figures 2 - 4 in section 3.4.

3.4 Elicitation task and procedure

Before the recording, participants indicated their demographic information (age, gender, language specifications, visual, hearing, or cognitive difficulties) in a short questionnaire. The stimuli presentation started with an introduction of the virtual character, Maria, who serves as a "conversation partner" throughout the recording. At the beginning of the task, the visual materials displayed the following information about Maria: she is a language learner of the participants' native language and she is currently learning about the colors and about giving instructions. Thus, participants were asked to help Maria by giving instructions mentioning the colored object she had to take. Additionally, participants were told that due to Maria being a language learner, she might make mistakes by choosing the objects of her favorite color, which is blue. In communicating with Maria, they were also asked to behave naturally by using their body to express themselves.

Participants were familiarized with the task by a practice phase prior to the experimental phase. The experimenter (Sánchez-Ramón for Catalan and Gregori for German) controlled the stimuli presentation. During the task, participants were prompted to produce sentences containing a target phrase with a Noun Phrase (NP) structure. In Catalan the structure is Article + Noun + Adjective, and in German it is Article + Adjective + Noun (e.g., Catalan: *Agafa [el gerro lila]^{NP}*; German: *Nimm [die lila Vase]^{NP}*; 'Take [the purple vase]^{NP}'). Words bearing stress on the first or the penultimate syllable were prompted in order to control for prosodic structure. All nouns and adjectives were frequently used and easily recognisable by the participants.

Each item started with a context slide, which reminded participants of their task and introduced the target item in its action context. Next, participants saw Marias' bag, containing the target item and potential competitor items. At this moment, participants were supposed to instruct Maria to take the correct object from the bag. The objects that appeared in the bag were manipulated in order to elicit the desired production. After Maria picked the correct object (or action), in all cases participants then gave feedback to Maria. Next, a white screen appeared that indicated carrying on with the next item.

In the information focus condition, only one object appeared in the bag (Figure 2), and the color of the item was not indicated in the context slide, to make the items coherent in the newness for production across all conditions.
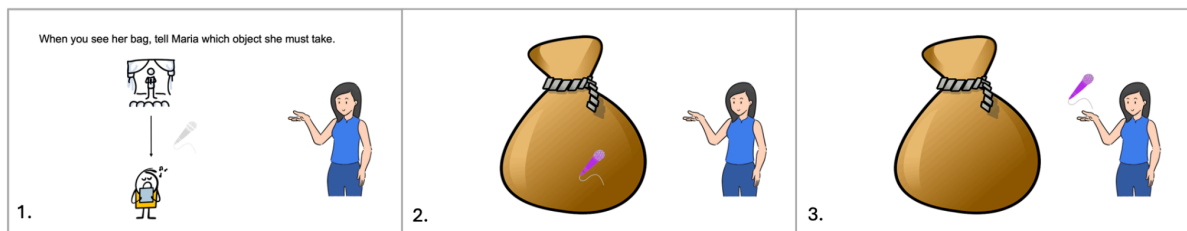
Figure 2: *Example trial of the study of an information focus item. 2.1 shows the context slide. 2.2 elicits the target production of the information focus item: "Maria, take the purple microphone." 2.3 shows Maria taking the correct object, prompting positive feedback.*

In the contrastive focus condition (Figure 3), a competitor object of a different color appeared in the bag. The corrective condition took place after an instruction in contrastive focus in which Maria made a mistake and took the wrong object (see Figure 3). Therefore, the participant had to correct her by repeating the instruction so that she could change her action.

Figure 3: *Example trial of the study of a contrastive and corrective focus item. 3.1 shows the context slide. 3.2 elicits the contrastive focus, showing two objects that contrast in color, target production: "Maria take the yellow glasses.". 3.3 elicits corrective focus, since the wrong object is picked, target production: "No, Maria, take the yellow glasses." 3.4 shows Maria taking the correct object, prompting positive feedback.*

Additionally, a contrastive filler condition contained two objects that contrast in color. It was included as an alternative condition in which Maria took one of the objects without making a mistake. Finally, in the action condition, participants first saw the same context slide, and, in a second display, two possible actions appeared. They had to decide which of the actions should be performed using the object (see Figure 4).
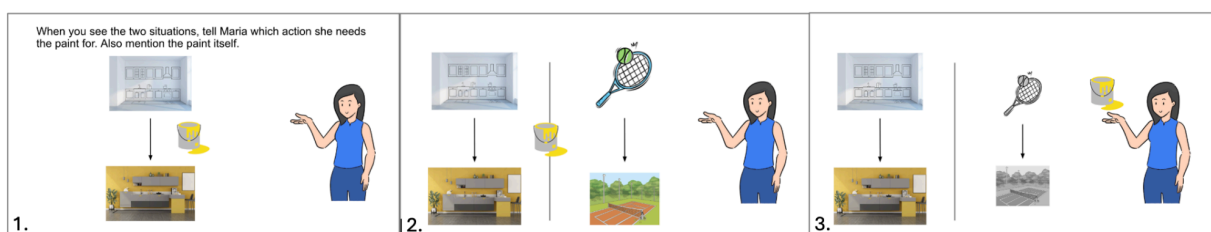

Figure 4: *Example trial of the study of an action condition item. 4.1 shows the context slide. 4.2 elicits the target production of the action condition: "Maria, take the yellow paint for painting the wall." 4.3 shows Maria taking the object for the correct action, prompting positive feedback.*

After finishing the recording, participants filled out a feedback questionnaire on the understandability of the task and their comfortability during the recording as well as the self-assessed naturalness of their use of speech and gestures using a scale from 1 to 7. Table 3 contains the mean self-indicated ratings on naturalness and comfortability of the participants. The results show that participants found the task easy to understand and felt comfortable in the recording environment. Additionally, they evaluated the naturalness of their own speech and gestures above average.

Table 3: *Overview of self-assessed task-understandability, comfortability and naturalness ratings of the participants. Ratings were given on a scale from 1(low) to 7(high).*

	Catalan	German
Task understandability	5.68	5.11
Comfortability	5.54	5.26
Naturalness of own speech	5.02	4.63
Naturalness of own gestures	4.42	4.69

The recording sessions took approximately 14 minutes on average (12.9 min in Catalan, 14.8 min in German), and the participants took approximately an additional 15 minutes for filling out the demographic and the feedback questionnaires.

4. Data annotation

The MultiFocus corpus contains manually annotated video and audio files. All prosodic annotations were done using Praat (Boersma & Weenink, 2024), and gestural annotations were done using ELAN (Sloetjes & Wittenburg, 2008, ELAN, 2024). The annotations will be uploaded to the language subcomponents in the form of .TextGrid and .eaf files after the completion of the current studies. The corpus contains orthographic, prosodic, and gestural annotations, described in sections 4.1, 4.2, and 4.3 respectively.

Sánchez-Ramón was responsible for the Catalan annotations, Gregori was responsible for the German annotations. All annotations were performed by a main annotator blind to the focus conditions and reliability annotations were done by the main annotators and two additional trained annotators. For detailed information on the annotation responsibilities per language, see the language subcomponents.

4.1 Orthographic and morphological annotation

In a first tier (see Figure 5), the complete utterances produced by the participants were identified in the recordings and named with a unique label. In a second tier, the utterances were orthographically annotated.

Next, target phrases (see "TP" tier in Figure 5), which generally contained both the noun and adjective (see tables 1 and 2 for examples of target phrases), were identified and segmented into words. Note that due to the semi-spontaneous speech, participants also produced target phrases which only contained the noun (Cat.: "Agafa les ulleres"; Ger.: "Nimm die Brille" *Take the glasses*), the adjective (Cat.: "Agafa les grogues; Ger.: Nimm die

gelben"; *Take the yellow (one)*), or a pronoun (Cat.: "Agafa les altres"; Ger.: "Nimm die andere" *Take the other (one)*).

Inside the target phrase, stressed syllables of the words, as well as their preceding and following syllables were annotated on the "Syll" tier (see Figure 5). Importantly, prosodic and gestural annotations were exclusively performed on the target phrase.

4.2 Prosodic annotation

Target phrases were annotated for prosody, specifically for pitch accent types and prosodic prominence. Annotations were performed in an audio-only setting without having access to the videos. Pitch accents were annotated in each target word following Cat_ToBI for Catalan (Prieto et al., 2015) and GToBI for German (Grice et al., 2005) on the "PA" tier. Prosodic prominence of the words within the target phrases was annotated following DIMA (Kügler et al., 2022) as a perceived holistic measure on a scale from level 0 (no prosodic prominence) to level 3 (extremely strong prosodic prominence). These annotations can be found on the "Prosodic_Prom" tier. A prosodic prominence rating of level 2 represents a typically strong pitch accent. Figure 5 displays the annotated prosodic tiers of an exemplary item in Praat.
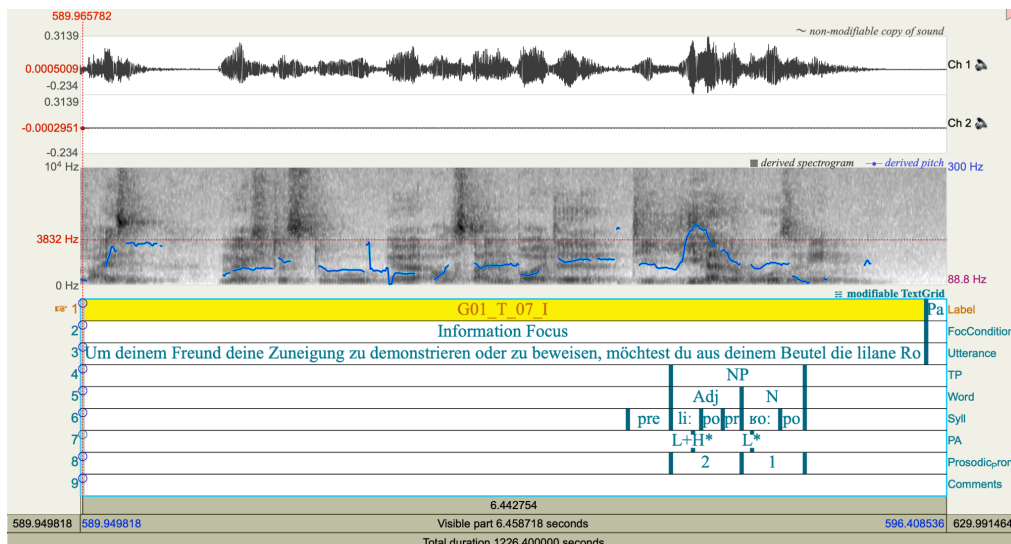

Figure 5: *Example of prosodic (and orthographic) annotations of a German item performed in Praat. The "PA" tier contains GToBI pitch accents, the "Prosodic_Prom" tier contains DIMA prominence ratings.*

4.3 Gesture annotation

Gesture annotations within the domain of the target phrases were done entirely without sound following M3D (Rohrer et al., 2023). The presence of gestures performed by hands, head, and eyebrows were annotated (see the tiers "Manual_GUnit", "Manual_GPhase", "Head_Movement_Type" and "EyebrowMovements" in Figure 6). Annotations on hand gestures included gesture phases. Each hand gesture stroke was also annotated for referentiality ("Semantic_ID" tier and its corresponding child tiers in Figure 6). Head movements were classified into types (nod, turn, tilt, protrusion, slide), and eyebrow movements were classified into types as well (both eyebrows raised, left eyebrow raised, right eyebrow raised, frowning eyebrows; not included in M3D). Gesture prominence is also annotated on a scale from 0 (no gesture prominence) to 3 (extremely strong gesture prominence) following M3D (Rohrer et al., 2023) as a perceived holistic measure of visual

cues. A gesture prominence rating of level 2 represents strong-prominence gestures. Each word within the target phrase received a rating of gesture prominence. Gesture prominence annotations can be found on the "Gestural_Prom" tier. For the purposes of this project, an additional level of "-1" was annotated in the words in which not any gesture was present. Figure 6 displays the annotated gestural tiers of an exemplary item in ELAN. All prosodic annotations were merged with the gestural annotations into one ELAN file afterwards.
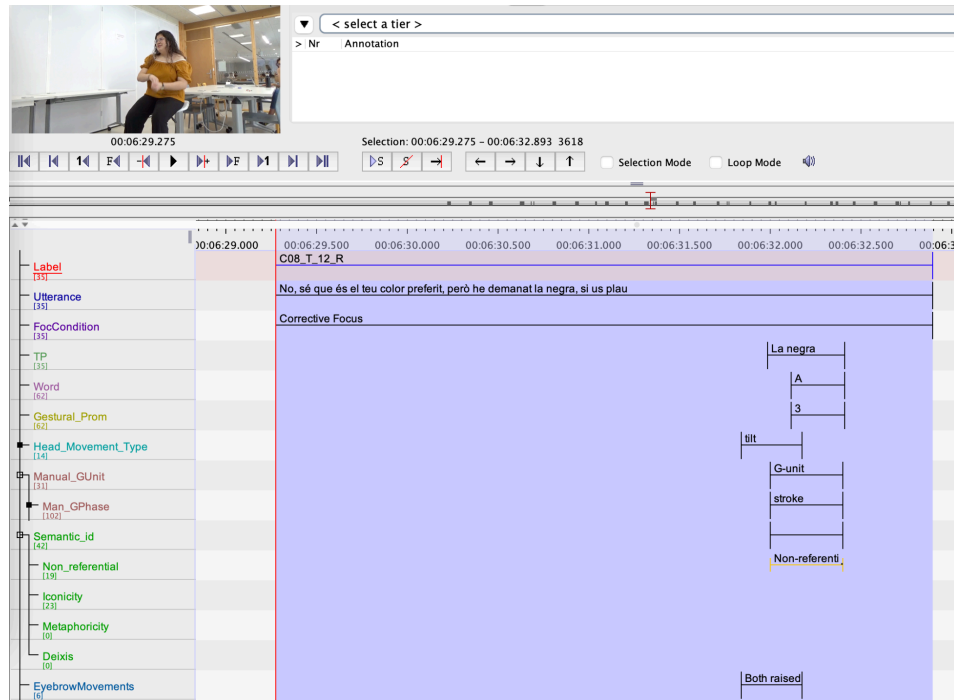

Figure 6: *Example of gesture (and orthographic) annotations of a Catalan item performed in ELAN. The "Manual_GUnit" and "Man_GPhase" tiers indicate hand gestures and their phases, the "Semantic_id" and its respective child tiers indicate referentiality. "Head_Movement_Type" and "EyebrowMovements" contain annotations of the respective articulator. The "Gestural_Prom" tier contains gesture prominence annotation, all tiers were annotated following M3D.*

5. Inter-rater reliability

In order to assure a reliable manual prominence annotation, inter-rater reliability of gesture presence, prosodic prominence and gesture prominence was assessed on 30% of the total database (11 participants per language). Those were performed by the main annotator and two other independent and trained native-speaker annotators. The secondary annotators were trained to annotate prosodic prominence with DIMA as well as gesture presence and gesture prominence using M3D in multiple small group sessions that included practice annotations in order to improve their understanding of the labelling systems. During the reliability assessment, doubtful cases were discussed among all raters (see language components for a detailed description of the reliability training).

Fleiss' Kappa (Fleiss, 1971; Fleiss et al., 2003) was used to calculate inter-rater reliability in gesture presence, as it is suitable for categorical data across multiple raters. Additionally, Gwet's AC2 (Gwet, 2008, 2014) were performed to calculate the inter-rater reliability on prosodic prominence ratings and gesture prominence ratings. In contrast to Kappa

measurements, these tests are able to not only grasp perfect agreement between multiple raters, but also their deviation in ordinal steps, integrating this into the accuracy measure. Results for gesture presence reveal substantial agreement for Catalan (z = 36.2, Kappa = 0.703, p < 0.0001) and a very good agreement for German (z = 36.9, Kappa = 0.818, p < 0.0001). Results for prosodic prominence reveal an almost perfect agreement in Catalan (coeff. val = 0.92122, conf. int. = (0.913,0.93), p < 0.0001) and similarly high agreement in German (coeff. val = 0.90463, conf. int. = (0.892, 0.918), p < 0.0001). For gestural prominence, the test also revealed an almost perfect inter-rater agreement in Catalan (coeff. val = 0.88513, conf. int. = (0.869, 0.901), p < 0.0001) as well as in German (coeff. val = 0.87167, conf. int. = (0.855, 0.888), p < 0.0001).

6. Corpus figures

The corpus contains 16.5 hours of recordings (7.5h in Catalan, 9h in German) and 2447 annotated productions (1223 sentences in Catalan and 1224 in German). See table 4 for a more detailed insight into the raw numbers of the corpus.

Table 4: *Important figures of the MultiFocus corpus, per language and in total*

Feature	Catalan	German	Total
Participants	35	35	70
Annotated Sentences	1223	1224	2447
Length	7.5 hours	9 hours	16.5 hours
Gestures in total	1447	2015	3462
Hand gestures (strokes)	573	508	1081
Head movements	716	1290	2006
Eyebrow movements	158	217	375
Pitch accents	2709	2284	4993

7. Organization of the corpus files

The OSF of the MultiFocus corpus is organized as follows: The general OSF page contains the corpus description, as well as materials (stimuli presentation, exemplary questionnaires on demographic data and feedback) in English. The corpus is linked to two subcomponents, one for each of the two languages. These subcomponents contain more detailed, language-specific information on stimuli, language structure and annotations as well as links to studies conducted on the respective languages using the MultiFocus corpus. Slides that were used for the recording for each language are also uploaded containing instructions and stimuli in the respective target language. In addition, in the subcomponents, we will upload .TextGrid (Praat) and .eaf (ELAN) annotation files of all recordings (but not the video and audio files, which are available upon request) when the current research endeavour is completed. Table 5 provides an overview of all files that will be uploaded to the OSF and where they will be saved.

Table 5: *Organization of the corpus files in the OSF*

MATERIAL	GENERAL MULTIFOCUS OSF osf.io/4m53q	LANGUAGE SUBCOMPONENTS CAT: https://osf.io/7z524/ GER: https://osf.io/86zeg/
Stimuli presentation - English	X	
Demographic questionnaire template - English	X	
Feedback questionnaire template - English	X	
Corpus subcomponent description - CAT/GER		X
List of prompted productions - CAT/GER		X
Stimuli presentation - CAT/GER		X
Demographic questionnaire template - CAT/GER		X
Feedback questionnaire template - CAT/GER		X
Annotated .TextTGrid files - CAT/GER		X
Annotated .eaf files - CAT/GER		X
Annotation trainings - CAT/GER		X

8. Usage restrictions

The collected data of the MultiFocus corpus contains sensitive personal information, participants are visible unmasked in the recorded videos. Therefore, the raw videos are not freely accessible on this OSF, but can be provided upon request, for the participants that agreed to their data being shared with other groups for scientific purposes (31 files for Catalan, 34 files for German). Please reach out to the first authors of the language subcomponents to request the video files and metadata. This OSF entry publishes materials relevant for setting up a follow-up experiment (including the experiments' stimuli slides). This corpus is available for scientific research purposes only.

9. Acknowledgements

10. References

Boersma, P. & Weenink, D. (2024). *Praat: doing phonetics by computer*. [Computer software]. Version 6.4.04. Retrieved from http://www.praat.org/

ELAN (2024). [Computer software]. Nijmegen: Max Planck Institute for Psycholinguistics, The Language Archive. Version 6.9. Retrieved from https://archive.mpi.nl/tla/elan

Esteve-Gibert, N., Lœvenbruck, H., Dohen, M., & D'Imperio, M. (2022). Pre-schoolers use head gestures rather than prosodic cues to highlight important information in speech. *Dev. Sci., 25(1)*, e13154.

Fleiss, J. L. (1971). Measuring nominal scale agreement among many raters. *Psychological Bulletin*, *76*, 378-382.

Fleiss, J. L., Levin, B., & Paik, M. C. (2003). *Statistical methods for rates and proportions* (3rd ed.). Hoboken, NJ: Wiley, 2003.

Grice, M., Baumann, S. & Benzmüller, R. (2005). German intonation in autosegmental-metrical phonology. Jun, S.-A. (Ed.), *Prosodic typology. The phonology of intonation and phrasing,* 55–83. Oxford University Press.

Gwet, K. L. (2008). Computing inter-rater reliability and its variance in the presence of high agreement. *The British Journal of Mathematical and Statistical Psychology, 61(1),* 29–48.

Gwet, K. L. (2014). *Handbook of inter-rater reliability: The definitive guide to measuring the extent of agreement among raters* (4th ed.) Gaithersburg, MD: Advanced Analytics.

Kügler, F., Baumann, S. & Röhr, C. (2022). Deutsche Intonation, Modellierung und Annotation (DIMA): Richtlinien zur prosodischen Annotation des Deutschen. Schwarze, C. & Grawunder, S. (eds.) *Transkription und Annotation gesprochener Sprache und multimodaler Interaktion: Konzepte, Probleme, Lösungen*, 23–54. Tübingen: Narr Francke Attempto Verlag.

Krifka, M. (2008). Basic notions of information structure. *Acta Linguist. Hung.,* 55(3–4), 243–276.

Prieto, P., Borràs-Comes, J., Cabré, T., Crespo-Sendra, V., Mascaró, I., Roseano, P., et al. (2015). Intonational phonology of Catalan and its dialectal varieties. In: Sónia Frota & Pilar Prieto (eds): *Intonation in Romance*. Oxford: OUP, 9–62.

Repp, S. (2014). Contrast: Dissecting an elusive information-structural notion and its role in grammar. In Féry, C. & Ishihara, S. (eds.), *Handbook of Information Structure*, OUP.

Rohrer, P. L., Tütüncübasi, U., Vilà-Giménez, I., Florit-Pons, J., Esteve-Gibert, N., Ren-Mitchell, A., Shattuck-Hufnagel, S., & Prieto, P. (2023). *The MultiModal MultiDimensional (M3D) labeling system.* https://doi.org/10.17605/osf.io/ankdx

Sloetjes, H., & Wittenburg, P. (2008). Annotation by category - ELAN and ISO DCR. *Proceedings of the 6th International Conference on Language Resources and Evaluation (LREC 2008)*.