

Annotation protocol for SOUND events within MAP(h)AS (Maneras de actuar, pensar, hablar y sentir en inglés y español/ Ways of behaving, thinking, speaking and feeling in English and Spanish. Funded by the Spanish Ministerio de Economía Industria and Competitividad MINECO, reference: FFI2017-86359-P).

The annotations were designed for a study conducted on three original narratives in English and three original narratives in Spanish (genres: thriller, romance, fantasy).

Step 1: Annotations with respect to domain instantiation

The first step was to identify expressions of SOUND events in the form of verb constructions (finite and non-finite). They may also be sequences consisting of support verbs such as *give*, *have* or *take* + a noun. Once the verbs had been identified, they were tagged with the general label TAGG for us to be able to easily retrieve the instances from the annotated corpora consisting of the entire texts. Due to copyright restrictions these annotations can neither be stored nor shared. TAGG was then followed by tags related to the type of verb construction in the first step. For instance,

- Overhead, they heardTAGG PERCEPTION&SOUND the excited flutter of [...]
- The trade winds blowing off the Coral Sea were warm and sweet, evocative reminders of faraway places that whistledTAGG PERCEPTION&SOUND through the rigging of the ketches and sloops [...]
- his voice carriedTAGG MOTION&SOUND loudlyTAGG ADV to all the surrounding tables

Step 2: Annotations with respect to perspectives

The second part of the analysis was concerned with the perspective from which the SOUND event was described in the texts, i.e., with focus on the Producer, Experiencer, or Phenomenon.

The table below summarizes the tags we use for the analysis of SOUND events.

TAG 1	TAG 2	TAG 3
Retrieval	Domain instantiation	Perspective
TAGG	PERCEPTION MOTION EMOTION-REACTION MANIPULATION CONSUMPTION COGNITION Support verb constructions*	Producer Experiencer Phenomenon

*The short TAG 2 for support verb constructions is AUX in the raw data txt files called SOUND-English_LC and SOUND-Spanish_LC.

The out-of-context raw data are given in the A-Tables in the List-of-tables pdf. This information includes the anchor verbs, domains, perspectives and number of occurrences in both the English and the Spanish data sets.

Tables	Domain instantiations	Perspectives	Languages
A-Table 4	PERCEPTION	Producer	English, Spanish
A-Table 5	PERCEPTION	Experiencer	English, Spanish
A-Table 6	PERCEPTION	Phenomenon	English, Spanish
A-Table 8	MOTION	Producer	English, Spanish
A-Table 9	MOTION	Experiencer	English, Spanish
A-Table 10	MOTION	Phenomenon	English, Spanish
A-Table 12	Support verb constructions	Producer, Experiencer, Phenomenon	English, Spanish
A-Table 13	MANIPULATION	Producer, Experiencer, Phenomenon	English, Spanish
A-Table 14	EMOTION-REACTION, CONSUMPTION, COGNITION	Producer, Experiencer, Phenomenon	English, Spanish