

## Normative DIS-TAG

The normative **DIS-TAG** created for the PRIN 2022 project *UNITE – Universally Inclusive Technologies to Practice English* is based on the **UCREL Semantic Analysis System (USAS)**, developed at UCREL, Lancaster, which assigns semantic categories to English words (Archer, Wilson and Rayson, 2002).

The aim was to develop tags marking references to both linguistic and extra-linguistic elements related to **mobility** and **the senses**. In some cases, tags were derived directly from existing USAS sub-categories; where no corresponding sub-category existed, new tags were developed.

Each tag follows the format **CATEGORY-SUBCATEGORY** → *Example terms*

### A. MOBILITY (MOB)

It refers to **manual actions, digital interactions, and body movements**

- **MOB-MAN (Physical Manipulation)** → *push, pull, grasp, grab, hold, move, agility, asleep, anti-bodies, hands, legs, belly button, birth mark, body clock, collar bone, clap*
- **MOB-DIG (Digital Interaction)** → *click, tap, scroll, swipe, press, select*
- **MOB-BODY (Body Movements)** → *walk, run, stand, sit, turn, enter, exit, doze off, fall to sleep, alight, approached, arising, arrived, bob, bow, breeze, bustling, clamber, collisions, depart, dispersing, emanates, exiting, fleeing, ambling along, barge past, beat a retreat, burst through, came by, climb on, come by, coming home, crash through*

### B. SENSES (SEN)

It refers to **perception and cognition**

- **SEN-VIS (Visual Perception)** → *watch, see, look, observe, focus, check, behold, beholding, blind, blinkers, eye, eyeball, eyesight, gaze, glance, glimpse, invisible, observational, optical, staring cast ... eyes over, eye up, keep ... eye open, kept tabs on, look through, lose sight of, naked eye*
- **SEN-AUD (Auditory Perception)** → *listen, hear, detect, overhear, acoustic, aloud, auditory, bangy, blaring, buzzy, cacophony, chimes, clangy, clanking, clonked, crackles, doink, echo, grunt, hearing impaired*
- **SEN-TAC (Tactile Perception)** → *touch, feel, grasp, handle, grip, hold, clasp, and seize*

### C. DISCOURSE TRIGGERS (DIS)

It refers to normative discourse structure like **imperatives, modal verbs or explanatory justifications**

- **DIS-IMP (Imperative Structures)** → *Click here, Look at this, Tap to continue*
- **DIS-MOD (Modal Verbs Expressing Obligation)** → *You must click, You should see, You need to tap*

- **DIS-EXP (Explanatory Justifications)** → *Since you clicked, Now that you see this, If you watch closely*
- **DIS-ABL → (Ability)** *able, unable, capable, incapable, competence, accomplish, ability, aptitude, astute, bent, brainy, canny, capable, cleverly, competence, cunning, de-skilled, dinny*

Annotation Format:

DIS-TAG annotations are inserted inline within the text.

Example:

<TAG tag="DIS-IMP"> <TAG tag="MOB-BODY"> stay </TAG> physically active </TAG>

### Corpus File Structure

Each file in the UNITE Corpus contains a full interaction between a learner and a chatbot system and they are named as follows:

annotated\_conversations/

- universityname\_01.txt
- universityname\_02.txt
- ...
- universityname\_70.txt

### Bibliography

Archer, D., Wilson, A., & Rayson, P. (2002). *Introduction to the USAS category system*. Lancaster University. [https://ucrel.lancs.ac.uk/usas/usas\\_guide.pdf](https://ucrel.lancs.ac.uk/usas/usas_guide.pdf)

### Recommended Citation

De Brasi, V., & Mongibello, A. (2026). Normative DIS-TAG\_UNITE [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.18960403>

DOI: <https://doi.org/10.5281/zenodo.18960403>

These Guidelines were developed within the framework of the PRIN 2022 project UNITE – Universally Inclusive Technologies to Practice English, funded by the Italian Ministry of University and Research under the European Union – NextGenerationEU program. Project Code 2022JB5KAL; CUP J53D23008070006.