|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Language family | Language | CDI Production | | | CHILDES | |
| N | Ages | ULs | Types | Tokens |
| Indo-European: Germanic | Danish | 6,112 | 8–36 | 601 | 4,955 | 195,658 |
| Dutch | 1,828 | 8–36 | 806 | 24,423 | 1,546,926 |
| English (American) | 10,939 | 8–30 | 670 | 72,704 | 17,899,676 |
| English (Australian) | 1,497 | 12–30 | 544 |
| English (British) | 22,589 | 12–47 | 522 |
| German | 1,181 | 18–30 | 538 | 63,832 | 3,745,393 |
| Norwegian | 7,358 | 8–36 | 658 | 10,428 | 224,402 |
| Swedish | 1,367 | 8–28 | 653 | 11,952 | 387,670 |
| Indo-European: Romance | Catalan | 1,436 | 8–30 | 571 | 9,363 | 196,637 |
| French (French) | 1,489 | 8–30 | 630 | 24,509 | 2,110,026 |
| French (Quebecois) | 1,760 | 8–30 | 625 |
| Italian | 1,400 | 7–36 | 624 | 10,424 | 294,050 |
| Portuguese (European) | 4,326 | 8–30 | 526 | 10,757 | 447,904 |
| Spanish (European) | 1,005 | 8–30 | 533 | 25,360 | 1,263,959 |
| Spanish (Mexican) | 1,985 | 8–30 | 619 |
| Indo-European: Slavic | Croatian | 627 | 8–30 | 614 | 12,831 | 226,081 |
| Czech | 493 | 16–30 | 477 | 16,636 | 193,573 |
| Russian | 1,805 | 8–36 | 653 | 4,184 | 38,591 |
| Sino-Tibetan | Cantonese | 1,208 | 16–30 | 688 | 11,281 | 1,483,524 |
| Mandarin (Beijing) | 1,938 | 12–30 | 807 | 18,532 | 1,740,448 |
| Mandarin (Taiwanese) | 2,654 | 8–36 | 626 |
| Uralic | Estonian | 1,235 | 16–30 | 570 | 28,803 | 560,396 |
| Hungarian | 363 | 16–30 | 662 | 19,321 | 224,062 |
| Japonic | Japanese | 846 | 8–36 | 569 | 261,159 | 703,143 |
| Koreanic | Korean | 2,422 | 8–36 | 599 | 45,736 | 395,720 |
| Semitic | Hebrew | 534 | 11–35 | 591 | 16,772 | 515,542 |
| Turkic | Turkish | 3,537 | 8–36 | 652 | 7,006 | 55,303 |

|  |  |  |  |
| --- | --- | --- | --- |
| Type | Item properties | Data origin | Processing resources |
| Phonological | Length in phonemes |  | eSpeak NG |
| Phonological neighbourhood |  | eSpeak NG |
| Morphological | Entropy over inflected forms | CHILDES | UDPipe |
| Number of morphological features | CHILDES | UDPipe |
| Number of morphemes | CHILDES | UniMorph |
| Syntactic | Mean length of utterance in words | CHILDES |  |
| Entropy over subcategorisation frames | CHILDES | UDPipe |
| Semantic | Concreteness | Adult norms  (Brysbaert et al., 2014) |  |
| Babiness | Adult norms  (Perry et al., 2015) |  |
| Distributional | Frequency | CHILDES |  |