



FP7 - Information and Communication Technologies

Collaborative Project

Dicta-Sign

**Sign Language Recognition, Generation and Modelling
with application in Deaf Communication**

Grant Agreement Number 231135

***Deliverable D6.3: Annotated Parallel Corpora for
BSL, DGS, GSL, LSF – Addendum***

Leading Partner: University of Hamburg (UHH)

Version: 2.0



Table of Contents

1	Synopsis	3
2	Corpus Content.....	3
2.1	Corpus Video Productions	3
2.2	Corpus Transcription.....	5
2.3	Metadata	6
3	The Dicta-Sign Web Portal	7
3.1	Access Levels.....	7
3.2	Layout.....	8
4	References	10

Appendix: The Dicta-Sign Web Portal – Task Descriptions and Content Tags

1 Synopsis

The aim of Workpackage 6 was to create the project language resources. Its objective was to produce reasonably sized annotated corpora and lexicons for the four European sign languages involved in the project (BSL, DGS, GSL, and LSF) and to make the corpus data available to other workpackages and the research community at large by providing annotated video data.

During the project lifetime, it was decided to make use of recent technological advances by including stereo cameras into the data collection process and to provide single-sign recordings in the same camera setup as extra training material for sign recognition.

Deliverable D6.3 comprises of these four subcorpora: The video footage and export files of the transcripts as well as video clips with annotation subtitles are made available on a server. A portal was created in order to present the material and allow convenient access to the data for a larger community. This addendum describes the corpus video production and postprocessing in the four countries as well as the transcription work. It introduces the Dicta-Sign web portal with its different access levels and the layout that allows for different access options to the corpus data.

2 Corpus Content

2.1 Corpus Video Productions

Recordings took place for all four sign languages involved in the project to collect data for Dicta-Sign's multilingual corpus. In preparation of the elicitation sessions the material described in D6.2 was adjusted for each country and task explanations were translated into the respective sign language. The elicitation sessions that took place in all four countries resulted in recorded material of at least 10.5 hours per language. Not taking into account data that had to be excluded for certain reasons (e.g. due to a withdrawal of the informant's consent) as well as parts like introduction, task explanations or feedback the corpus video inventory now consists of at least 8 hours signing time from 14 to 16 different informants per language (see table below for details).

	DGS	LSF	GSL	BSL
Signing Time (hh:min)	09:07	08:20	10:07	08:12
No. of Sessions	7	8	8	8
No. of Informants	14	16	16	16

Table 1: Dicta-Sign video inventory

The complex studio setup that was decided to be used for Dicta-Sign's data collection consisted of seven cameras, two of them stereo cameras (Hanke et al. 2010a). The different camera perspectives (front, side and bird's eye view) were to help annotators interpret the signing. The additional stereo cameras provide footage that allows image analysis to reconstruct 3D information and help automatic processing.

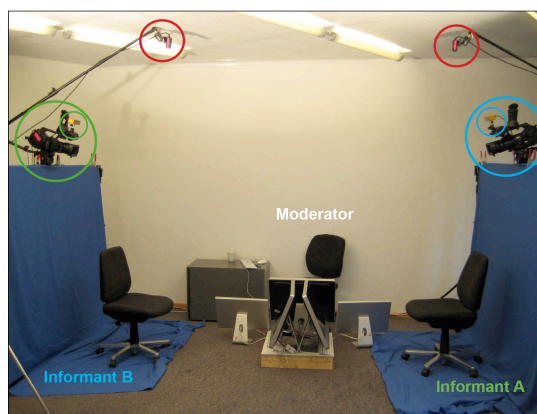


Figure 1: The Dicta-Sign studio at UHH

A variety of postprocessing steps were needed before annotation work could start, most importantly providing backup data, compression of the video files as well as precise frame-by-frame synchronisation. iLex was further developed to handle multi-camera recordings (Hanke et al. 2010b). It now allows synchronising video clips of the different camera perspectives by adjusting the time code information (see Figure 2). In the transcript window (Figure 3) different layout options may be defined to adjust arrangement and size of the different perspectives according to the annotators' individual needs. As an extra step in postprocessing, UHH decided to convert their video clips, recorded at a frame rate of 25 frames per second interlaced, to 50fps by deinterlacing the video footage. This allows for more precise segmentation of the individual signs.¹



Figure 2: Synchronising camera perspectives in iLex

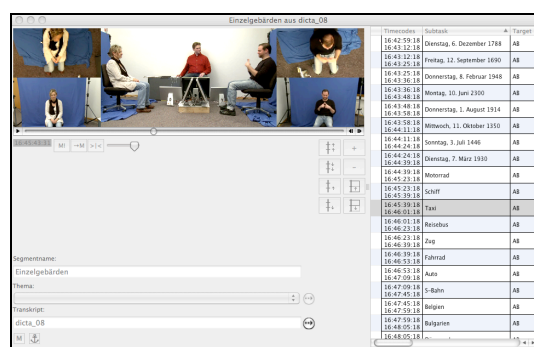


Figure 3: iLex transcript window with multicamera layout

Dicta-Sign's basic lexicon list, initially planned as a bootstrapping phase for corpus annotation, was later decided to be used as a shared list of 1000 concepts for the four sign languages. In order to provide reliable video footage for WP2 at an early stage it was further decided to additionally record the 1000 lemmas in a studio reproduction. For GSL and DGS the recordings comprise of 5-10 times the same sign of one signer (for GSL) or two signers (for DGS) which served as training material for sign recognition. For LSF and BSL lemmas were recorded one time each. Postprocessing was done as for the corpus data. However, an extra step was needed for the DGS data due to recording errors in one of the cameras that caused a loss of frames and made synchronisation of the video clips impossible. The iLex synchronisation function was therefore extended in a way to allow inserting frames where needed.

¹ With the goal of improving the avatars' performance DGS videos with 50fps were also used to gain more information on the exact timing of individual parameters in signs (cf. Hanke et al. 2011).

2.2 Corpus Transcription

Corpus transcription work has been carried out for all four sign languages within the project. In a first step, new iLex databases had to be created for BSL, GSL and LSF. For DGS it was possible to take advantage of the already existing database being used at UHH. The video data were then integrated into iLex and transcripts were produced for all tasks. At UHH, where Session Director had been used to run the elicitation sessions (cf. Hanke et al. 2010a), it was possible to provide automatic tagging specifying start and end of the individual tasks and subtasks using time information from the Session Director log files.

Due to the delays in corpus collection the annotation only started in the second year of the project. In order to minimise the impact of these delays on other workpackages, especially WP4, this workpackage determined the order in which the basic annotation was carried out. It was agreed between the project partners to annotate the “Airport Task” (task 4) as a first step (now available for at least 5 sessions per language), followed by annotation of as many different tasks as possible for each partner (see Table 2 for details). The annotation process for all sign languages included segmentation, lemmatisation and form description. With regard to segmentation of the continuous signing it was decided to treat transitional movements between individual signs not as part of either sign (i.e. there are gaps between two signs during which the articulators move from the end of one sign to the beginning of the next). This approach, though more time consuming than segmenting once in the middle of a transition, offers advantages for the subsequent processing within Dicta-Sign: Firstly it results in the fact that a token tag then represents only that part of the signal that is described by HamNoSys. Secondly, variation between tokens is much lower than if the transition would be part of the sign.

After segmentation the individual signs could be lemmatised, i.e. unique glosses were assigned by means of type-token matching in the growing database, and a HamNoSys description was added on type level.

As a further step to enrich the data annotation, individual project partners conducted extra annotation work: For the DGS data mouth patterns were annotated by assigning either written German words to represent mouthings or “MG” as a preliminary tag for mouth gestures. Furthermore, in addition to the annotation of HamNoSys on type level form deviations between types and the respective tokens were tagged as such in order to provide more reliable training data for image processing. The GSL data provide a tagging of clause boundaries, and for LSF pointing, buoys and depicting signs have been annotated using categories close to those proposed in the Auslan annotation guideline provided by Trevor Johnston², and for pointing and depicting signs some indication on the use of the signing space have been added. Furthermore, English translation is provided for parts of the German, Greek, and French subcorpora as well as a French translation for the majority of the LSF data.

The elicitation sessions for the Dicta-Sign corpus had been planned in a way to provide more data than initially needed, and therefore not all videos that are now available could be fully annotated in the framework of the project. In order to allow access to all corpus data, it was therefore decided to provide a coarse tagging by assigning content tags that reflect the topics covered by the individual informants (see Table 2).

² <http://www.auslan.org.au/video/upload/attachments/AuslanCorpusAnnotationGuidelines30November2011.pdf>

Task	Lemmatisation / Content Tags (No. of Sessions)			
	DGS	LSF	GSL	BSL
1	1 / 6	5 / 8	8 / 8	0 / 8
2	1 / 6	5 / 8	8 / 8	0 / 8
3	1 / 6	5 / 8	8 / 8	0 / 8
4	7 / 7	5 / 8	8 / 8	8 / 0
5	0 / 6	5 / 7	0 / 8	0 / 8
6	1 / 6	5 / 8	8 / 8	0 / 8
7	1 / 6	5 / 3	8 / 8	0 / 8
8a	2 / 6	5 / 8	8 / 8	8 / 0
8b	2 / 6	5 / 8	7 / 8	0 / 8
9	1 / 6	8 / 0	0 / 8	0 / 8
10	0 / 5	0 / 0	0 / 0	0 / 0

Table 2: Annotated Video Data

The Dicta-Sign multilingual corpus is not a parallel corpus in the classical sense as the “oral” nature of sign language as well as the risk of influences from written majority languages do not allow for such an approach. Instead, the aim was to elicit semantically close answers without predetermining the choice of vocabulary and grammar. In order to allow finding video sections in the corpus with comparable content across individual informants and languages, consistent naming of the content tags had to be assured. For the majority of the tasks, e.g. “At the airport”, many topics covered by informants are directly related to the elicitation material and defining the topics across the different languages was comparably easy. For other tasks, as e.g. “Public transportation”, a mapping of information was needed: Station names vary across languages and many different routes were described by the individual informants. In order to allow for consistent content tagging codes were applied to the different routes and pictures were produced accordingly (see Appendix). The topics found in the Dicta-Sign corpus and their occurrences across the different languages are being evaluated and results will be reported at the '5th Workshop on the Representation and Processing of Sign Languages' (LREC2012).³

2.3 Metadata

Personal metadata was collected from the informants by means of questionnaires based on the IMDI standard with sign language-specific extensions as defined in Crasborn/Hanke 2003. As that set covers a variety of purposes for metadata (e.g. to support language acquisition studies), but does not explicitly define subparts, Dicta-Sign defined a subset that seemed suitable for the kind of study conducted here and also minimised the questionnaire filling effort for the informants.

When the data collection was planned, it was evident that metadata would be collected in a finer granularity than appropriate for publication. It was therefore communicated to the informants (and agreed by them through the Informed Consents) that only coarsened data would be made available without specifying exactly what that means. We had hoped that projects focussing on metadata (esp. CLARIN and

³ Matthes et al. (in preparation): Dicta-Sign – Building a Multilingual Sign Language Corpus.

NaLiDa) would meanwhile define standards how to specify such coarsenings so that the sign language corpus community could fill out the details on sign language data. Unfortunately, although there has been a lot of communication between projects and the Dicta-Sign data were selected as a case study for CLARIN-D, this is not yet the case. Instead, we defined two levels of coarsening for different publicity levels (see below on portal structure) of informant data. E.g. the informant's date of birth is converted to the age in years for restricted access and age range (e.g. 41-50) for public access.

For the time being, data are made available in IMDI session file format. Through cooperation with NaLiDa/CLARIN-D, these data will be converted to the upcoming CMDI component structure.

The coarsenings applied were deliberately chosen to be restrictive. In cases where users of the Dicta-Sign data would require more specific data, it was agreed that they should contact the individual Dicta-Sign partners responsible for the subcorpus concerned. The procedure for verifying the scientific need and contacting the informant to agree is then identical to the one for requests to contact informants for extra data collections.

3 The Dicta-Sign Web Portal

A web portal was developed to allow access to the Dicta-Sign multilingual corpus for public use as well as research purposes. It can be accessed from the Dicta-Sign website: <http://www.dictasign.eu/Main/Portal>.

3.1 Access Levels

There are three accessibility levels to the corpus data:

- Publicly available on the project site: Metadata and “appetisers” to the video and transcript data. Information on elicitation materials is at the same level of detail as previously published (Matthes et al. 2010).
- Restricted access for researchers: Elicitation materials available for researchers' own purposes as well as to fully understand the data collection. Video and transcript data as available in a standard format (H.264 for video, ELAN and iLex export format for transcripts).
- The third level is not available online, but requires arrangements between the researcher interested and the individual partner owning the data. This includes higher-resolution less compressed video and stereo data and more refined metadata.

Additionally the web portal includes contact forms for researchers who request higher-resolution and stereo data from individual partners or ask to contact informants to be given access to more detailed metadata or to suggest additional data collection.

3.2 Layout

The Dicta-Sign web portal offers different approaches to access the corpus data:

- **By Language & Informants:** For every sign language the available recording sessions (i.e. pairs of informants) are listed. Via the sessions all tasks performed by the respective informants as well as informant metadata information can be accessed.

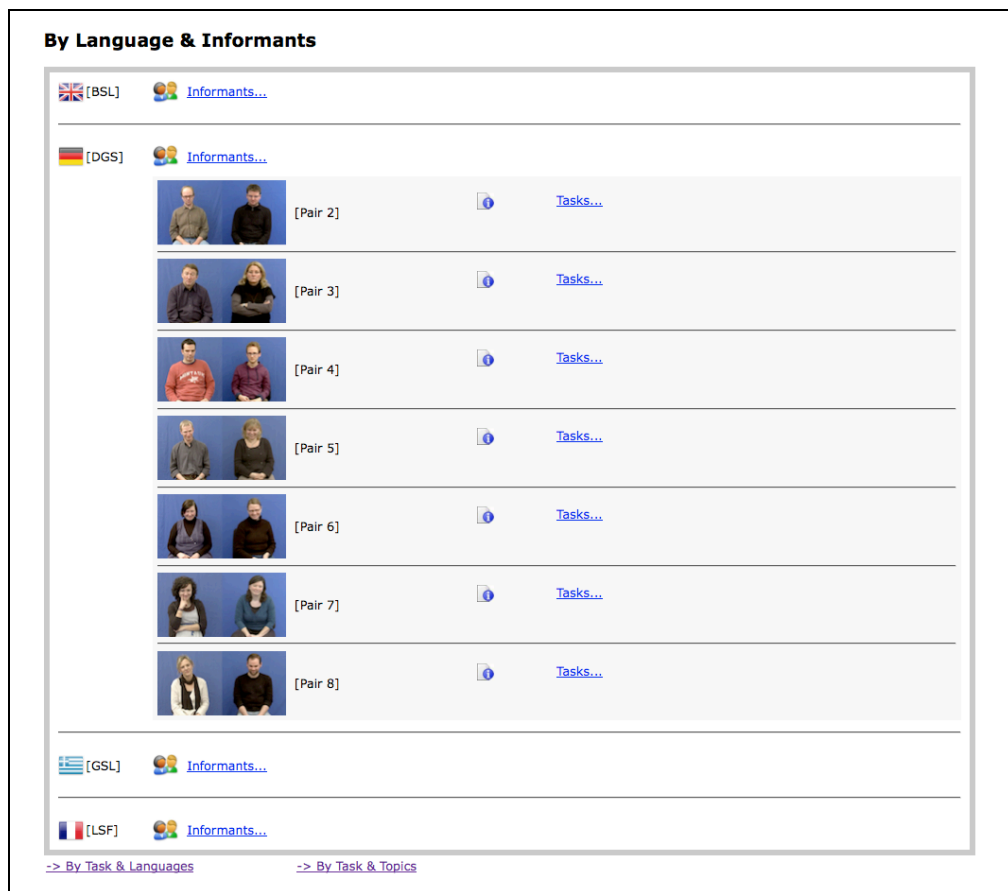


Figure 4

- **By Task & Languages:** For each task that the informants were asked to perform a short description as well as the elicitation material is provided (cf. Appendix). Grouped by languages, all data-by-task items for a certain task are listed and can be accessed. In addition, the content tags defined for each task are presented.

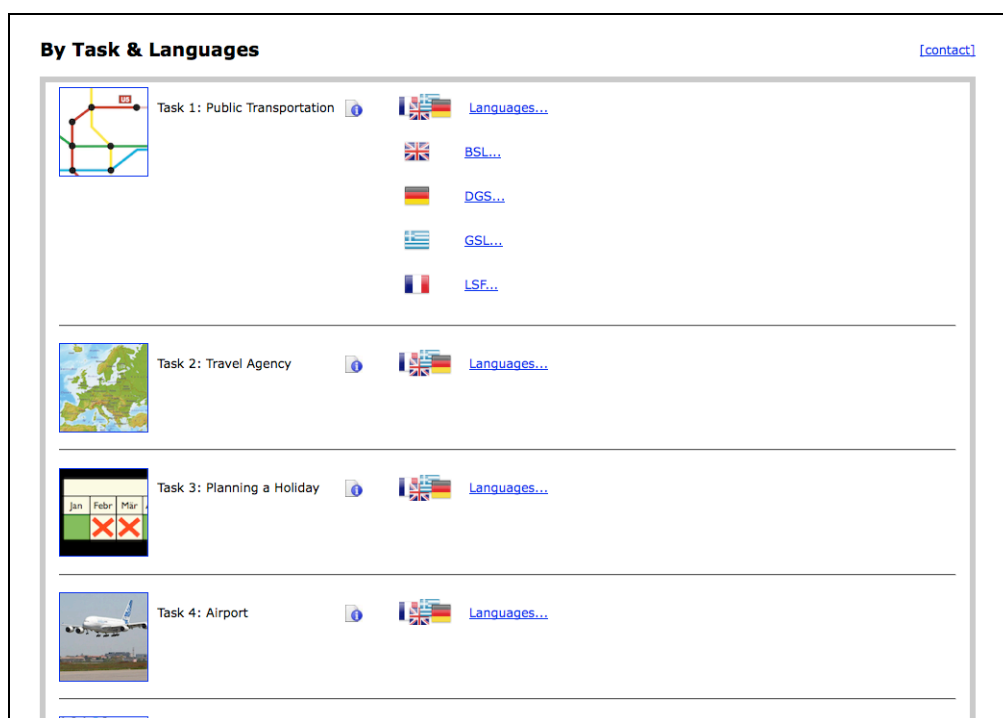


Figure 5

- **By Task & Topics:** This approach makes use of the topics identified as part of the annotation process. By listing all content tags of the individual tasks it allows access to comparable data across individual informants and languages (cf. Appendix).

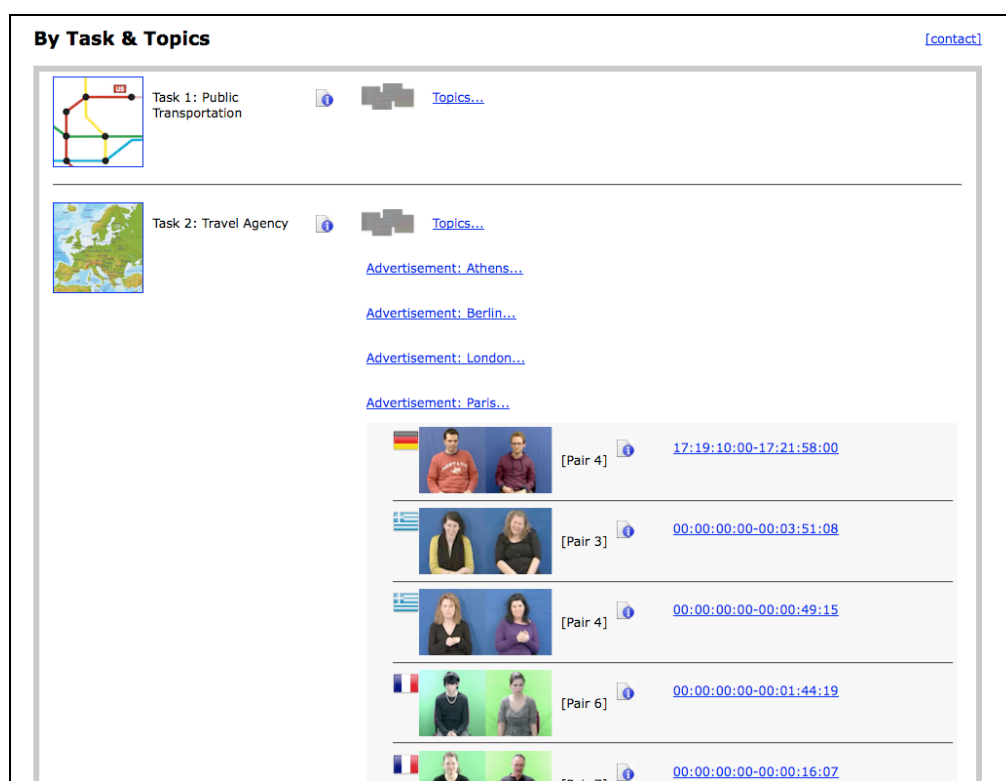
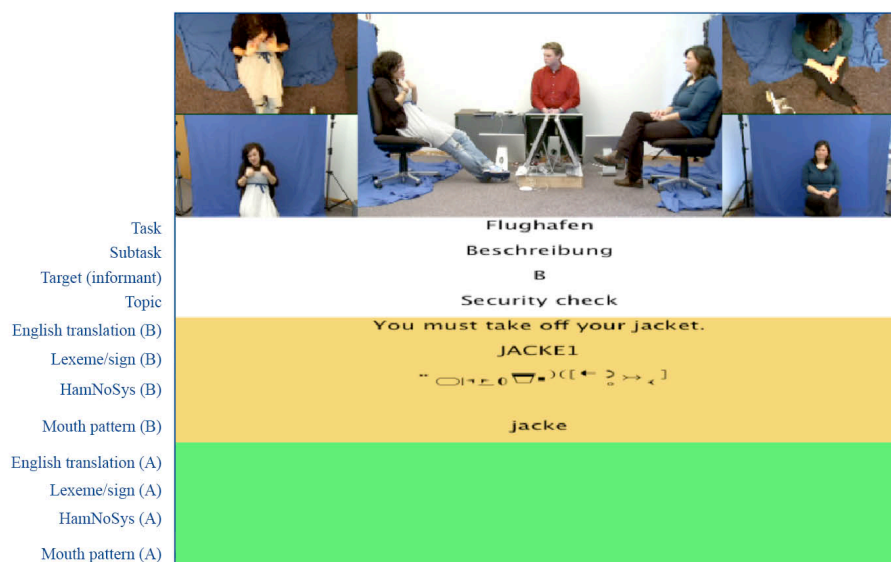


Figure 6

The corpus video and transcription data linked to the portal via the different access options are made available as videos with and without subtitles (cf. Figure 7), in iLex export format as well as ELAN format. The elicitation material including task explanations in the respective sign language is provided as Keynote or PowerPoint documents. The portal also contains a section for access to the 10x1000 lemmas that were recorded based on Dicta-Sign's basic lexicon list.

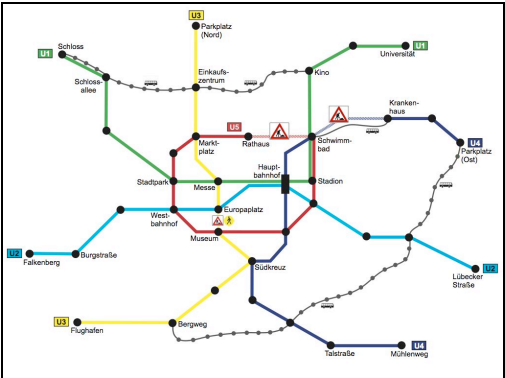


Task	Subtask	Target (informant)	Topic
English translation (B)	Lexeme/sign (B)	HamNoSys (B)	Mouth pattern (B)
English translation (A)	Lexeme/sign (A)	HamNoSys (A)	Mouth pattern (A)

Flughafen
Beschreibung
B
Security check
You must take off your jacket.
JACKE1
" O H E 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046 1047 1048 1049 1050 1051 1052 1053 1054 1055 1056 1057 1058 1059 1060 1061 1062 1063 1064 1065 1066 1067 1068 1069 1070 1071 1072 1073 1074 1075 1076 1077 1078 1079 1080 1081 1082 1083 1084 1085 1086 1087 1088 1089 1090 1091 1092 1093 1094 1095 1096 1097 1098 1099 1100 1101 1102 1103 1104 1105 1106 1107 1108 1109 1110 1111 1112 1113 1114 1115 1116 1117 1118 1119 1120 1121 1122 1123 1124 1125 1126 1127 1128 1129 1130 1131 1132 1133 1134 1135 1136 1137 1138 1139 1140 1141 1142 1143 1144 1145 1146 1147 1148 1149 1150 1151 1152 1153 1154 1155 1156 1157 1158 1159 1160 1161 1162 1163 1164 1165 1166 1167 1168 1169 1170 1171 1172 1173 1174 1175 1176 1177 1178 1179 1180 1181 1182 1183 1184 1185 1186 1187 1188 1189 1190 1191 1192 1193 1194 1195 1196 1197 1198 1199 1200 1201 1202 1203 1204 1205 1206 1207 1208 1209 1210 1211 1212 1213 1214 1215 1216 1217 1218 1219 1220 1221 1222 1223 1224 1225 1226 1227 1228 1229 1230 1231 1232 1233 1234 1235 1236 1237 1238 1239 1240 1241 1242 1243 1244 1245 1246 1247 1248 1249 1250 1251 1252 1253 1254 1255 1256 1257 1258 1259 1260 1261 1262 1263 1264 1265 1266 1267 1268 1269 1270 1271 1272 1273 1274 1275 1276 1277 1278 1279 1280 1281 1282 1283 1284 1285 1286 1287 1288 1289 1290 1291 1292 1293 1294 1295 1296 1297 1298 1299 1300 1301 1302 1303 1304 1305 1306 1307 1308 1309 1310 1311 1312 1313 1314 1315 1316 1317 1318 1319 1320 1321 1322 1323 1324 1325 1326 1327 1328 1329 1330 1331 1332 1333 1334 1335 1336 1337 1338 1339 1340 1341 1342 1343 1344 1345 1346 1347 1348 1349 1350 1351 1352 1353 1354 1355 1356 1357 1358 1359 1360 1361 1362 1363 1364 1365 1366 1367 1368 1369 1370 1371 1372 1373 1374 1375 1376 1377 1378 1379 1380 1381 1382 1383 1384 1385 1386 1387 1388 1389 1390 1391 1392 1393 1394 1395 1396 1397 1398 1399 1400 1401 1402 1403 1404 1405 1406 1407 1408 1409 1410 1411 1412 1413 1414 1415 1416 1417 1418 1419 1420 1421 1422 1423 1424 1425 1426 1427 1428 1429 1430 1431 1432 1433 1434 1435 1436 1437 1438 1439 1440 1441 1442 1443 1444 1445 1446 1447 1448 1449 1450 1451 1452 1453 1454 1455 1456 1457 1458 1459 1460 1461 1462 1463 1464 1465 1466 1467 1468 1469 1470 1471 1472 1473 1474 1475 1476 1477 1478 1479 1480 1481 1482 1483 1484 1485 1486 1487 1488 1489 1490 1491 1492 1493 1494 1495 1496 1497 1498 1499 1500 1501 1502 1503 1504 1505 1506 1507 1508 1509 1510 1511 1512 1513 1514 1515 1516 1517 1518 1519 1520 1521 1522 1523 1524 1525 1526 1527 1528 1529 1530 1531 1532 1533 1534 1535 1536 1537 1538 1539 1540 1541 1542 1543 1544 1545 1546 1547 1548 1549 1550 1551 1552 1553 1554 1555 1556 1557 1558 1559 1560 1561 1562 1563 1564 1565 1566 1567 1568 1569 1570 1571 1572 1573 1574 1575 1576 1577 1578 1579 1580 1581 1582 1583 1584 1585 1586 1587 1588 1589 1590 1591 1592 1593 1594 1595 1596 1597 1598 1599 1600 1601 1602 1603 1604 1605 1606 1607 1608 1609 1610 1611 1612 1613 1614 1615 1616 1617 1618 1619 1620 1621 1622 1623 1624 1625 1626 1627 1628 1629 1630 1631 1632 1633 1634 1635 1636 1637 1638 1639 1640 1641 1642 1643 1644 1645 1646 1647 1648 1649 1650 1651 1652 1653 1654 1655 1656 1657 1658 1659 1660 1661 1662 1663 1664 1665 1666 1667 1668 1669 1670 1671 1672 1673 1674 1675 1676 1677 1678 1679 1680 1681 1682 1683 1684 1685 1686 1687 1688 1689 1690 1691 1692 1693 1694 1695 1696 1697 1698 1699 1700 1701 1702 1703 1704 1705 1706 1707 1708 1709 1710 1711 1712 1713 1714 1715 1716 1717 1718 1719 1720 1721 1722 1723 1724 1725 1726 1727 1728 1729 1730 1731 1732 1733 1734 1735 1736 1737 1738 1739 1740 1741 1742 1743 1744 1745 1746 1747 1748 1749 1750 1751 1752 1753 1754 1755 1756 1757 1758 1759 1760 1761 1762 1763 1764 1765 1766 1767 1768 1769 1770 1771 1772 1773 1774 1775 1776 1777 1778 1779 1780 1781 1782 1783 1784 1785 1786 1787 1788 1789 1790 1791 1792 1793 1794 1795 1796 1797 1798 1799 1800 1801 1802 1803 1804 1805 1806 1807 1808 1809 1810 1811 1812 1813 1814 1815 1816 1817 1818 1819 1820 1821 1822 1823 1824 1825 1826 1827 1828 1829 1830 1831 1832 1833 1834 1835 1836 1837 1838 1839 1840 1841 1842 1843 1844 1845 1846 1847 1848 1849 1850 1851 1852 1853 1854 1855 1856 1857 1858 1859 1860 1861 1862 1863 1864 1865 1866 1867 1868 1869 1870 1871 1872 1873 1874 1875 1876 1877 1878 1879 1880 1881 1882 1883 1884 1885 1886 1887 1888 1889 1890 1891 1892 1893 1894 1895 1896 1897 1898 1899 1900 1901 1902 1903 1904 1905 1906 1907 1908 1909 1910 1911 1912 1913 1914 1915 1916 1917 1918 1919 1920 1921 1922 1923 1924 1925 1926 1927 1928 1929 1930 1931 1932 1933 1934 1935 1936 1937 1938 1939 1940 1941 1942 1943 1944 1945 1946 1947 1948 1949 1950 1951 1952 1953 1954 1955 1956 1957 1958 1959 1960 1961 1962 1963 1964 1965 1966 1967 1968 1969 1970 1971 1972 1973 1974 1975 1976 1977 1978 1979 1980 1981 1982 1983 1984 1985 1986 1987 1988 1989 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019 2020 2021 2022 2023 2024 2025 2026 2027 2028 2029 2030 2031 2032 2033 2034 2035 2036 2037 2038 2039 2040 2041 2042 2043 2044 2045 2046 2047 2048 2049 2050 2051 2052 2053 2054 2055 2056 2057 2058 2059 2060 2061 2062 2063 2064 2065 2066 2067 2068 2069 2070 2071 2072 2073 2074 2075 2076 2077 2078 2079 2080 2081 2082 2083 2084 2085 2086 2087 2088 2089 2090 2091 2092 2093 2094 2095 2096 2097 2098 2099 2100 2101 2102 2103 2104 2105 2106 2107 2108 2109 2110 2111 2112 2113 2114 2115 2116 2117 2118 2119 2120 2121 2122 2123 2124 2125 2126 2127 2128 2129 2130 2131 2132 2133 2134 2135 2136 2137 2138 2139 2140 2141 2142 2143 2144 2145 2146 2147 2148 2149 2150 2151 2152 2153 2154 2155 2156 2157 2158 2159 2160 2161 2162 2163 2164 2165 2166 2167 2168 2169 2170 2171 2172 2173 2174 2175 2176 2177 2178 2179 2180 2181 2182 2183 2184 2185 2186 2187 2188 2189 2190 2191 2192 2193 2194 2195 2196 2197 2198 2199 2200 2201 2202 2203 2204 2205 2206 2207 2208 2209 2210 2211 2212 2213 2214 2215 2216 2217 2218 2219 2220 2221 2222 2223 2224 2225 2226 2227 2228 2229 2230 2231 2232 2233 2234 2235 2236 2237 2238 2239 2240 2241 2242 2243 2244 2245 2246 2247 2248 2249 2250 2251 2252 2253 2254 2255 2256 2257 2258 2259 2260 2261 2262 2263 2264 2265 2266 2267 2268 2269 2270 2271 2272 2273 2274 2275 2276 2277 2278 2279 2280 2281 2282 2283 2284 2285 2286 2287 2288 2289 2290 2291 2292 2293 2294 2295 2296 2297 2298 2299 2300 2301 2302 2303 2304 2305 2306 2307 2308 2309 2310 2311 2312 2313 2314 2315 2316 2317 2318 2319 2320 2321 2322 2323 2324 2325 2326 2327 2328 2329 2330 2331 2332 2333 2334 2335 2336 2337 2338 2339 2340 2341 2342 2343 2344 2345 2346 2347 2348 2349 2350 2351 2352 2353 2354 2355 2356 2357 2358 2359 2360 2361 2362 2363 2364 2365 2366 2367 2368 2369 2370 2371 2372 2373 2374 2375 2376 2377 2378 2379 2380 2381 2382 2383 2384 2385 2386 2387 2388 2389 2390 2391 2392 2393 2394 2395 2396 2397 2398 2399 2400 2401 2402 2403 2404 2405 2406 2407 2408 2409 2410 2411 2412 2413 2414 2415 2416 2417 2418 2419 2420 2421 2422 2423 2424 2425 2426 2427 2428 2429 2430 2431 2432 2433 2434 2435 2436 2437 2438 2439 2440 2441 2442 2443 2444 2445 2446 2447 2448 2449 2450 2451 2452 2453 2454 2455 2456 2457 2458 2459 2460 2461 2462 2463 2464 2465 2466 2467 2468 2469 2470 2471 2472 2473 2474 2475 2476 2477 2478 2479 2480 2481 2482 2483 2484 2485 2486 2487 2488 2489 2490 2491 2492 2493 2494 2495 2496 2497 2498 2499 2500 2501 2502 2503 2504 2505 2506 2507 2508 2509 2510 2511 2512 2513 2514 2515 2516 2517 2518 2519 2520 2521 2522 2523 2524 2525 2526 2527 2528 2529 2530 2531 2532 2533 2534 2535 2536 2537 2538 2539 2540 2541 2542 2543 2544 2545 2546 2547 2548 2549 2550 2551 2552 2553 2554 2555 2556 2557 2558 2559 2560 2561 2562 2563 2564 2565 2566 2567 2568 2569 2570 2571 2572 2573 2574 2575 2576 2577 2578 2579 2580 2581 2582 2583 2584 2585 2586 2587 2588 2589

Appendix: The Dicta-Sign Web Portal – Task Descriptions and Content Tags

Task 1: Public Transportation Map

Category:	Route description
Main target:	Informant A/B
Description:	<p>The informants are asked to explain how to get from a certain place to another using public transportation. A map is provided to both of them displaying different means of public transport (underground, connecting busses, closed lines, possibility of walking) and stations (airport, town hall, train station, market, etc.). The task includes five subtasks where different stations are given as departure and destination points, indicated as flags on the map. Each of the informants is asked to suggest one possible route per subtask.</p> <p><i>German transportation map for DGS informants:</i></p> 
Topics:	<p>R1.1a: Einkaufszentrum (U3) - Europaplatz (walking) - Museum (U3) - Flughafen</p> <p>R1.1b: Einkaufszentrum (U3) - Flughafen <i>[without walking, not possible]</i></p> <p>R1.2: Einkaufszentrum (U3) - Europaplatz (U2) - Westbahnhof (U5) - Museum (U3) - Flughafen</p> <p>R1.3: Einkaufszentrum (U3) - Südkreuz (U4) - <i>unnamed station</i> (Bus) - Bergweg (U3) - Flughafen <i>[without walking, not possible]</i></p> <p>R1.4: Einkaufszentrum (U3) - Marktplatz (U5) - Museum (U3) - Flughafen</p> <p>R1.5: Einkaufszentrum (Bus) - Schlossallee (U1) - Stadtpark (U5) - Museum (U3) - Flughafen</p> <p>R1.6: Einkaufszentrum (U3) - Messe (U1) - Hauptbahnhof (U4) - Südkreuz (U3) - Flughafen</p> <p>R1.7: Einkaufszentrum (U3) - Messe (U1) - Stadion (U5) - Museum (U3) - Flughafen</p> <p>R1.8: Einkaufszentrum (Bus) - Kino (U1) - Hauptbahnhof (U4) - Südkreuz (U3) - Flughafen</p> <p>R1.9: Einkaufszentrum (U3) - Messe (U1) - Stadtpark (U5) - Museum (U3) - Flughafen</p>

R1.10:	Einkaufszentrum (Bus) - Kino (U1) - Stadion (U5) - Museum (U3) - Flughafen
R1.11:	Einkaufszentrum (U3) - Europaplatz (U2) - Hauptbahnhof (U4) - Südkreuz (U3) - Flughafen
R1.12:	Einkaufszentrum (Bus) - Kino (U1) - Schwimmbad (U5) - <i>unnamed station</i> (U2) - Hauptbahnhof (U4) - Südkreuz (U3) - Flughafen
R2.1:	Talstraße (U4) - Südkreuz (U3) - Museum (U5) - Westbahnhof (U2) - Burgstraße
R2.2:	Talstraße (U4) - Hauptbahnhof (U2) - Burgstraße
R2.3:	Talstraße (U4) - Hauptbahnhof (U1) - Stadtpark (U5) - Westbahnhof (U2) - Burgstraße
R2.4:	Talstraße (U4) - Südkreuz (U3) - Museum (walking) - Europaplatz (U2) - Burgstraße
R2.5:	Talstraße (U4) - Hauptbahnhof (U1) - Messe (U3) - Europaplatz (U2) - Burgstraße
R2.6:	Talstraße (U4) - Schwimmbad (U1) - Kino (Bus) - Schlossallee (U1) - Stadtpark (U5) - Westbahnhof (U2) - Burgstraße
R2.7:	Talstraße (U4) - Schwimmbad (U5) - <i>unnamed station</i> (U2) - Burgstraße
R2.8:	Talstraße (U4) - <i>unnamed station</i> (U5) - Westbahnhof (U2) - Burgstraße
R2.9:	Talstraße (U4) - Schwimmbad (U1) - Kino (Bus) - Einkaufszentrum (U3) - Europaplatz (U2) - Burgstraße
R2.10:	Talstraße (U4) - <i>unnamed station</i> (Bus) - (U2) - Europaplatz (walking) - Museum (U5) - Westbahnhof (U2) - Burgstraße
R3.1:	Parkplatz (Nord) (U3) - Einkaufszentrum (Bus) - Schloss
R3.2:	Parkplatz (Nord) (U3) - Marktplatz (U5) - Stadtpark (U1) - Schloss
R3.3:	Parkplatz (Nord) (U3) - Messe (U1) - Schloss
R3.4:	Parkplatz (Nord) (U3) - Einkaufszentrum (Bus) - Schlossallee (U1) - Schloss
R3.5:	Parkplatz (Nord) (U3) - Europaplatz (U2) - Westbahnhof (U5) - Stadtpark (U1) - Schloss
R3.6:	Parkplatz (Nord) (U3) - Europaplatz (walking) - Museum (U5) - Stadtpark (U1) - Schloss
R3.7:	Parkplatz (Nord) (walking) - Schloss
R4.1:	Flughafen (U3) - Bergweg (Bus) - Parkplatz (Ost)
R4.2:	Flughafen (U3) - Südkreuz (U4) - Schwimmbad (Bus) - Krankenhaus (U4) - Parkplatz (Ost)
R4.3:	Flughafen (U3) - Südkreuz (U4) - Hauptbahnhof (U7) - (Bus) - Parkplatz (Ost)
R4.4:	Flughafen (U3) - Museum (walking) - Europaplatz (U2) - Hauptbahnhof (U4) - Schwimmbad (Bus) - Krankenhaus (U4) - Parkplatz (Ost)
R4.5:	Flughafen (U3) - Südkreuz (U4) - <i>unnamed station</i> (U5) - <i>unnamed station</i> (U2) - <i>unnamed station</i> (Bus) - Parkplatz (Ost)
R4.6:	Flughafen (U3) - Südkreuz (U4) - <i>unnamed station</i> (U5) - Schwimmbad (Bus) - Krankenhaus (U4) - Parkplatz (Ost)
R4.7:	Flughafen (U3) - Museum (U5) - Schwimmbad (Bus) - Krankenhaus (U4) - Parkplatz (Ost)
R4.8:	Flughafen (U3) - Museum (U5) - <i>unnamed station</i> (U4) - Hauptbahnhof (U2) - <i>unnamed station</i> (Bus) - Parkplatz (Ost)
R4.9:	Flughafen (U3) - Messe (U1) - Schwimmbad (Bus) - Krankenhaus (U4) - Parkplatz (Ost)

R5.1:	Rathaus (walking, taxi, or bike) - Schwimmbad
R5.2:	Rathaus (U5) - Schwimmbad (<i>counter clockwise</i>)
R5.3:	Rathaus (U5) - Marktplatz (U3) - Einkaufszentrum (Bus) - Kino (U1) - Schwimmbad
R5.4:	Rathaus (U5) - Marktplatz (U3) - Museum (U5) - Schwimmbad (<i>without walking, not possible</i>)
R5.5:	Rathaus (U5) - Museum (U3) - Südkreuz (U4) - Schwimmbad
R5.6:	Rathaus (U5) - Stadtpark (U1) - Messe (U3) - Europaplatz (walking) - Museum (U5) - Schwimmbad
R5.7:	Rathaus (U5) - Marktplatz (U3) - Messe (U1) - Schwimmbad
R5.8:	Rathaus (U5) - Stadtpark (U1) - Schwimmbad
R5.9:	Rathaus (U5) - Marktplatz (U3) - Messe (U1) - Hauptbahnhof (U4) - Schwimmbad
Further discussion: Route	
Further discussion: Travelling by bus	
Further discussion: Other	

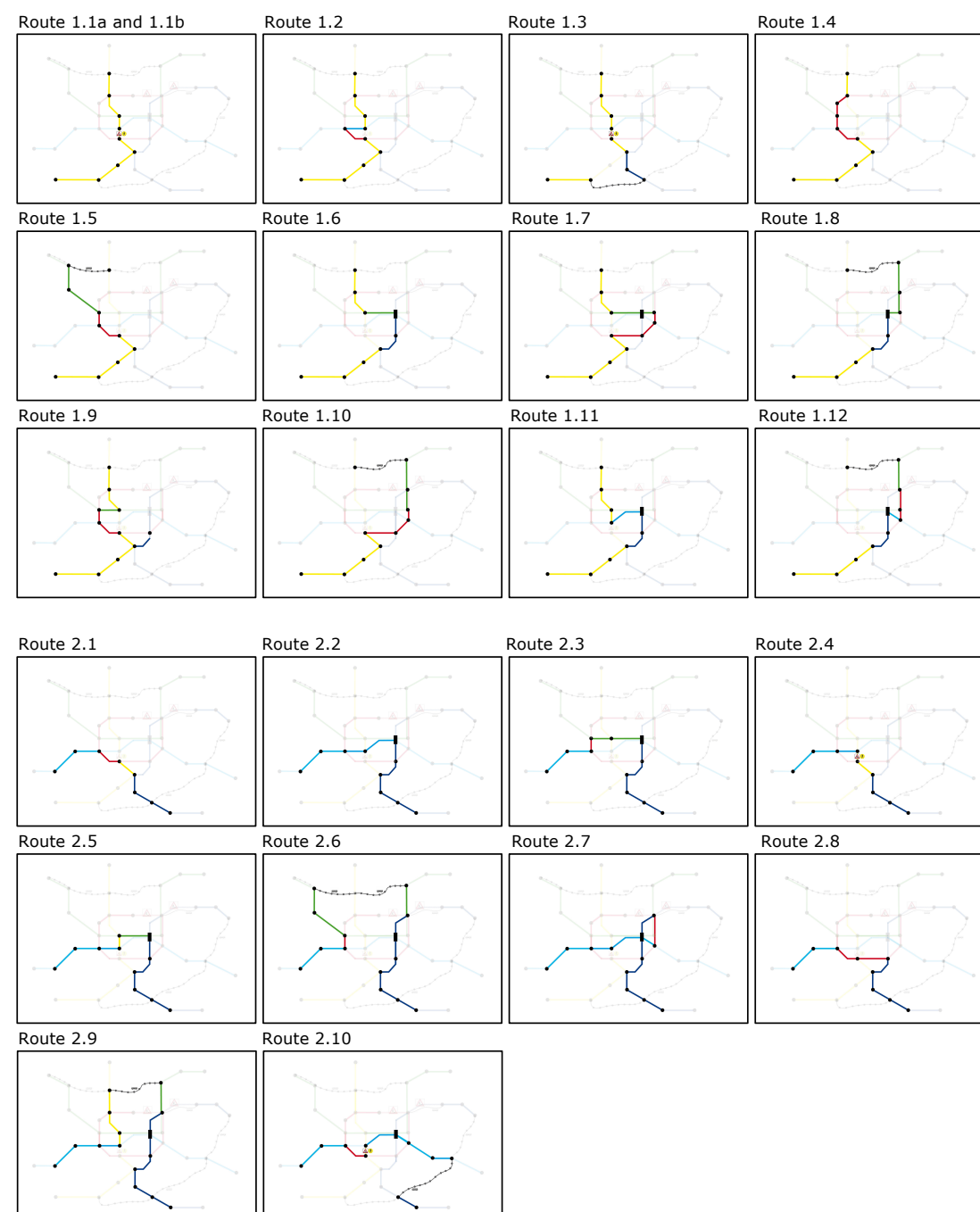
The names of the stations had to be adjusted depending on the target language. Codes were therefore assigned to the different routes found in the data to allow for consistent tagging across languages. Additionally, pictures of the routes were produced to ease the mapping.

Mapping table for station names:

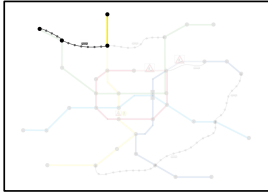
English	German	Greek	French
Castle	Schloss	Ολυμπιακό Στάδιο	Château
The Avenue	Schlossallee	Σταθμός Ειρήνη	Avenue du château
Town Park	Stadtpark	Ομόνοια	Jardin Public
Exhibition Centre	Messe	Πανεπιστήμιο	Salon d'exposition
Central Railway Station	Hauptbahnhof	Σύνταγμα	Gare centrale
Stadium	Stadion	Λεωφόρος Αλεξάνδρας	Stade
Swimming Pool	Schwimmbad	Μοναστηράκι	Piscine
Cinema	Kino	Πατήσια	Cinéma
University	Universität	Πλατεία Ζωγράφου	Université
Windsor Road	Falkenberg	Εθνικό Θέατρο	Pars d'attraction
Station Road	Burgstraße	Νοσοκομείο Ευαγγελισμός	Rue de la Tozr
Railway Station West	Westbahnhof	Λόφος Λυκαβητού	Gare de l'Ouest
Europe Square	Europaplatz	Εθνικός Κήπος	Place de l'Europe
London Street	Lübecker Straße	Νέος Κόσμος	Port maritime
Airport	Flughafen	Νέο Φάληρο	Aéroport
Green Lane	Bergweg	Άλιμος	Rue de la montagne
Southern Junction	Südkreuz	Λόφος Φιλοπάππου	Croix du sud
Museum	Museum	Μουσείο Ακρόπολης	Musée
Marketplace	Marktplatz	Πλατεία Ψυρρή	Place du marché

Shopping centre	Einkaufszentrum	Σταθμός Λαρίσης	Centre commercial
Car park (North)	Parkplatz (Nord)	Γκάζι	Parking Nord
Park Road	Mühlenweg	Αεροδρόμιο Ελ. Βενιζέλος	Chemin du Moulin
Kings Road	Talstraße	Εθνική Άμυνα	Rue de la vallee
Hospital	Krankenhaus	Δικαστήρια	Hôpital
Car park (East)	Parkplatz (Ost)	Αμπελόκηποι	Parking Est
Town Hall	Rathaus	Πλατεία Εξαρχείων	Mairie

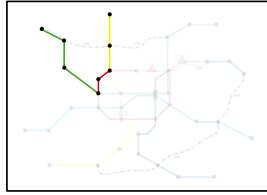
Routes described by the informants:



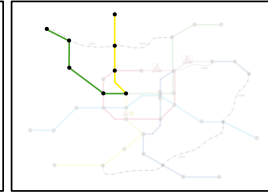
Route 3.1



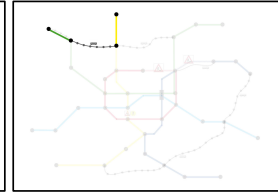
Route 3.2



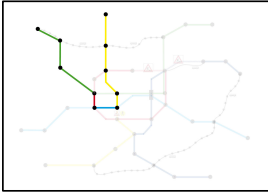
Route 3.3



Route 3.4



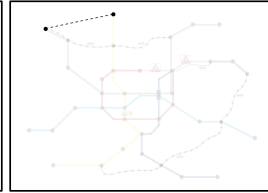
Route 3.5



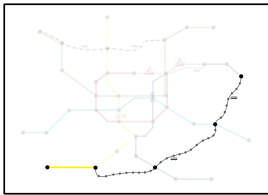
Route 3.6



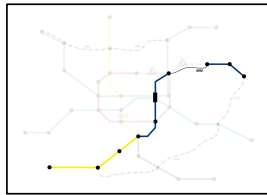
Route 3.7



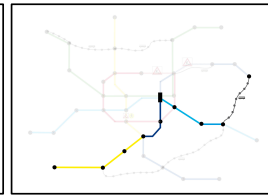
Route 4.1



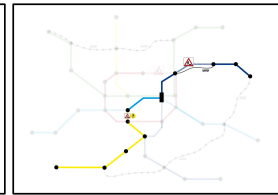
Route 4.2



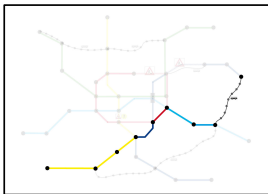
Route 4.3



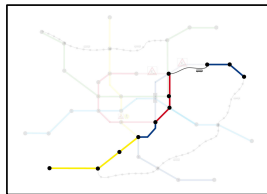
Route 4.4



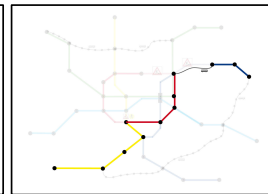
Route 4.5



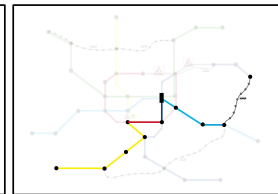
Route 4.6



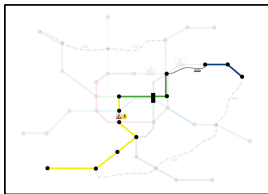
Route 4.7



Route 4.8



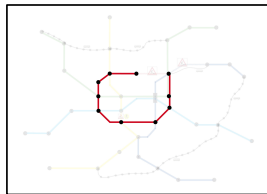
Route 4.9



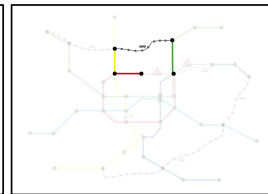
Route 5.1



Route 5.2



Route 5.3



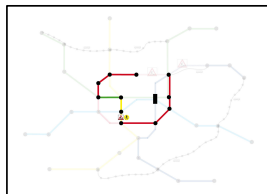
Route 5.4



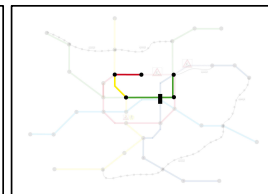
Route 5.5



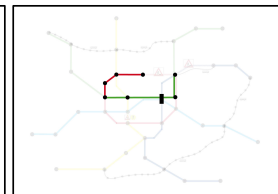
Route 5.6



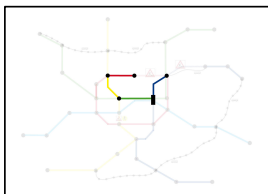
Route 5.7




Route 5.8



Route 5.9



Task 2: Travel Agency

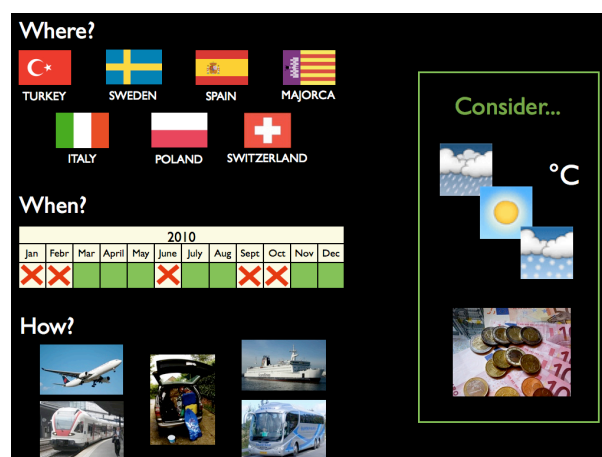
Category:	Description of places and activities
Main target:	Informant A
Description:	<p>The imagined setting for this task is a travel agency. The role of one of the informants is that of a staff member at the travel agency, while the other takes over the role of a customer who wishes to go on holiday but does not know yet where to go to. The staff member is asked to suggest two different destinations to the customer by describing/advertising these places. Afterwards the customer is asked to explain briefly which of the destinations she/he prefers and why.</p> <p>The two destinations, predefined by the elicitation material, are a combination of one capital city and a more general place and vary across the different sessions. For each destination a range of pictures are shown meant to provide ideas of what to talk about.</p> <p><i>Destinations to be described:</i></p> <div data-bbox="416 880 1340 1344">  </div>
Topics:	<p>Advertisement: Athens Advertisement: Berlin Advertisement: London Advertisement: Paris Advertisement: At the beach Advertisement: In the mountains Decision making</p>

Task 3: Planning a Holiday

Category:	Discussion and negotiation
Main target:	Informant A/B
Description:	The informants are instructed to plan a holiday together. They are asked to negotiate the destination, the time period for the holiday and the

means of transport. They should also take into account further aspects relevant for their decision. As a basis of their discussion a picture card is shown to each of the informants displaying flags of several countries/regions, different means of transport, as well as a calendar where certain time periods are blocked (only the calendar is different for the two informants). Additionally certain aspects are shown that should also be taken into account (weather, temperature, costs).


English version for BSL (informant A):



Topics:
Destination
Time
Transportation
Accommodation
Booking
Other

Task 4: At the Airport

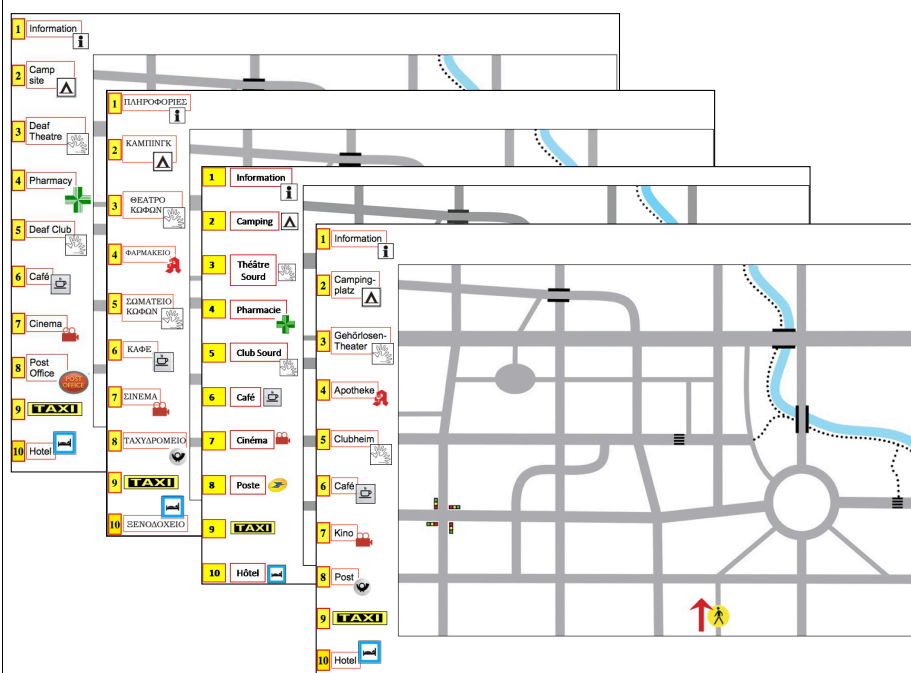
Category:	Description of places and activities
Main target:	Informant B
Description:	The topic of this task is the situation at an airport and the procedures taking place when travelling by plane. The informant is asked to describe the situation as if the other informant has never travelled by plane before. Pictures displaying different aspects as checking in, security issues, boarding, baggage claim and passport control are shown to the informant in chronological order.

	<p><i>Version for DGS (informant B):</i></p> 
Topics:	<p> Ticket purchase Preparation of the trip Transport to the airport Check-In Information board Security check Getting to the waiting room At the gate Airplane fuel On board: Seating On board: Safety instructions On board: Food and drinks On board: Amusement activities Baggage claim Customs and passport control Other </p>

Task 5: City Map

Category:	Route description
Main target:	Informant A
Description:	<p>Based on a map provided, one of the informants has to describe a walk through the city and to name several landmarks. The informant serving as the information giver has a route marked on her/his map as well as several landmarks. For the information follower a map is provided on paper which has a numbered list of the same landmarks on the side of the map but no route. Taking off from the starting point marked on both maps the information giver's task is to describe the route displayed on the map. Whenever a landmark is passed along the route, the information giver is asked to tell the information follower what it is and where exactly it is located. The latter is supposed to follow the route descriptions and note on her/his map where the landmarks are to be found.</p>

Maps for information followers provided on paper:



Topics:

Start - Information
Information - Camping site
Camping site - Deaf theatre
Deaf theatre - Pharmacy
Pharmacy - Deaf club
Deaf club - Café
Café - Cinema
Cinema - Post office
Post office - Taxi
Taxi - Hotel

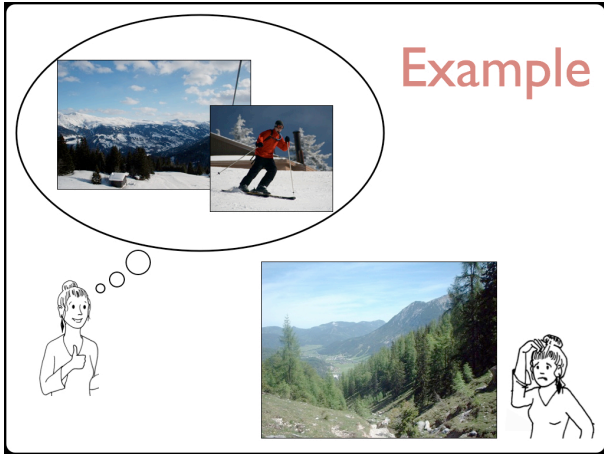
The names of the landmarks had to be adjusted depending on the target language. A translation is given in the mapping table.

Mapping table for landmarks:

English	German	Greek	French
Start	Start	Αρχή	Départ
Information	Information	Πληροφορίες	Information
Camping site	Campingplatz	Κατασκήνωση	Camping
Deaf theatre	Gehörlosentheater	Θέατρο Κωφών	Théâtre Sourd
Pharmacy	Apotheke	Φαρμακείο	Pharmacie
Deaf club	Clubheim	Σωματείο Κωφών	Club Sourd
Café	Café	Καφετέρια	Café

Cinema	Kino	Σινεμά	Cinéma
Post office	Post	Ταχυδρομείο	Poste
Taxi	Taxi	Στάση Ταξί	Taxi
Hotel	Hotel	Ξενοδοχείο	Hôtel

Task 6: Expectation & Reality


Category:	Narration
Main target:	Informant A/B
Description:	<p>The informants are asked to tell short stories based on picture cards showing two opposed occurrences of a certain situation (somebody's expectations and the actual situation). The following situations are included in the task:</p> <ul style="list-style-type: none"> • Skiing holiday / no snow (example) • Comfortable hotel room / tiny room with small bed • Summer holiday / flight cancelled • Visit to a nice museum/ overcrowded museum • Plenty of food in a restaurant / plate with little food • Garden party with BBQ / bad weather • Sunset at the beach / traffic jam, arriving in the dark <p>The informants are free to tell a true story (where something similar has happened to themselves) or make one up.</p> <p><i>Example shown to both informants in the task explanation:</i></p> 
Topics:	<p>Hotel room: Expectation Hotel room: Reality Flight: Expectation Flight: Reality Museum: Expectation</p>

Task 7: Travel Then & Now

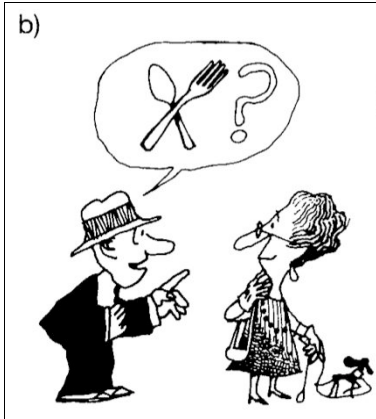

xi

	Destination_then Destination_now Lifestyle_then Lifestyle_now Technology_then Technology_now
--	---

Task 8a: Retelling of a Signed Story

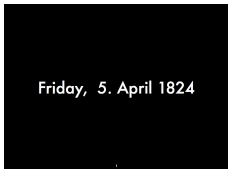


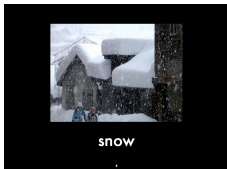
Category:	Narration
Main target:	Informant A
Description:	<p>A short clip of a signed story is shown where a Deaf person tells about what happened during his last holiday trip to Italy: His travel group arrived late at the hotel, and after a drink at the bar he went straight to bed and slept all night. The next morning the others tell him that the fire alarm went off during the night and how they tried to wake him up. The informant is asked to retell the story.</p> <p><i>Screenshot of the DGS video clip:</i></p> 
Topics:	Group journey to Italy At the hotel Going to bed In the morning Reaction of travel companions What happened: Rescue vans and evacuation What happened: Door locked What happened: No keys What happened: Banging the door What happened: Firefighter with axe What happened: False alarm Conclusion / relief

Task 8b: Retelling of a Picture Story

Category:	Narration
Main target:	Informant B
Description:	<p>A picture story by Quino (Lavado 1991) is used for this task, in which a woman explains to a tourist how to get to a certain restaurant (including walking, taking a taxi and a plane). The informant is asked to look at the story picture by picture and tell it to the other informant afterwards.</p> <p><i>Pictures b) and e) from the picture story :</i></p> <div style="display: flex; justify-content: space-around; align-items: center;">   </div>
Topics:	<ul style="list-style-type: none"> a) Introduction b) Looking for a restaurant c) Route description: Get a taxi d) Route description: Go to the airport e) Route description: Take a plane f) Route description: Landing in Paris g) Route description: Maxim's h) Thank you i) Wondering what to do j) Outdoor food stall k) Hiding Additional information

Task 9: Isolated Signs

Category:	Denomination
Main target:	Informant A/B

Description:	<p>Single pictures and, where needed for clarification, written words are used as stimuli. The informants are asked in turns to give isolated signs for the concepts shown. They are also encouraged to add to the other person's answers whenever they know a different sign for a certain concept.</p> <p>The task covers the following areas: Dates, vehicles, European countries, and weather conditions.</p> <p><i>Examples from the English version:</i></p> <div data-bbox="416 544 644 712">  </div> <div data-bbox="649 544 877 712">  </div> <div data-bbox="882 544 1110 712">  </div> <div data-bbox="1115 544 1343 712">  </div>
Topics:	<p>Monday, 2. November 2013 Friday, 5. April 1824 Wednesday, 9. January 1570 Saturday, 4. May 2005 Tuesday, 6. December 1788 Friday, 12. September 1690 Thursday, 8. February 1948 Monday, 10. June 2300 Thursday, 1. August 1914 Wednesday, 11. October 1350 Sunday, 3. July 1446 Tuesday, 7. March 1930</p> <p>Motorbike Boat Taxi Coach Train Bicycle Car Suburban railway</p> <p>Belgium Bulgaria Denmark Germany Estonia Finland France Greece Ireland Italy Croatia Latvia Lithuania</p>



	Luxembourg Malta Macedonia Netherlands Austria Poland Portugal Romania Sweden Switzerland Slovakia Slovenia Spain Czech Republic Turkey Hungary Great Britain Cyprus
	Sun Rain Storm Thunderstorm Fog Hail Snow

Task 10: Personal Experience

Category:	Narration
Main target:	Informant A/B
Description:	<p>Extra task: The informants are asked to tell a short story about a personal travel experience (e.g. their worst or best holiday).</p> <p><i>No material is needed for the task.</i></p>
Topics:	Egypt Majorca Croatia Italy Holland India Finland Denmark France Black Forest Other



Task 10 serves as an extra task that was only conducted if time was left at the end of the sessions and was therefore only performed by a small number of informants. As the content of the data resulting from this task was not expected to be similar across individual informants it was not tagged in detail. Instead a short summary of the informants' stories was produced. The topics listed above indicate the countries the stories take place in.

References

Lavado, Joaquin (1991): Guten Appetit! Oldenburg: Lappan.