**Supplementary information**

1) Detailed information on participants’ proficiency across three trials of data collection

Trial-English (Trial 1): A total of 59 participants were recruited and tested at Adam Mickiewicz University in Poznań, (Romance Languages Institute), but the analysis involved data from 55 participants. Students were recruited from four different levels of study: C1 - a small group consisted of 4 students (Portuguese Studies). Two of them had some exposure to Spanish, whereas the other two were fourth-year students of Spanish Philology; C2 - 12 second-year students of Spanish philology, after approximately 500h of practical Spanish classes. However, there might have been some differences in language levels, as some of them had started studying Spanish before they entered the university; C3 - 21 first-year students of Spanish philology, after approx. 200h of practical Spanish classes (similarly to the previous group, mixed levels may have been involved); C4 - 18 third-year students of Spanish philology, after more than 700h of Spanish instruction; they represented the highest level of proficiency which corresponded to B2/C1 level of the CEFR.

Trial-Polish (Trial 2): A total of 37 participants were recruited from the same university. Students were recruited from three different levels of study: C5 - a small group consisted of 4 fifth-year students of Spanish philology; C2 – 17 second-year students of Spanish philology, after approx. 500h of practical Spanish instruction; C3 – 16 third-year students of Spanish philology, after approx. 700h of practical Spanish classes.

Trial-Spanish (Trial 3): A total of 43 participants were recruited from the same university at two different levels of study: H1 – 22 first-year students of Spanish philology, after approx. 200h of practical Spanish classes; H2 – 21 second-year students of Spanish philology students, after approx. 500h of practical Spanish classes. See Table 1 for an additional comparison (ANOVAs and t-test) of students’ proficiency across levels of study:

Table 1.

L1-Polish students of Spanish and their proficiency scores across three trials

|  |  |  |  |
| --- | --- | --- | --- |
| Group/year of study | No. of participants (n) | Proficiency | Post-hoc tests |
|  |  | M (SD) |  |
| TRIAL-ENGLISH | n = 55 |  | *F*(3,51) = 13.44, *p* < .001 |
| C1 (4th year) | 4 | .70 (.26) |  |
| C2 (2nd year) | 12 | .83 (.06) |  |
| C3 (1st year) | 21 | .62 (.20) | C3 < C2\* |
| C4 (3rd year) | 18 | .90 (.05) | C3 < C4\* |
| TRIAL-POLISH | n = 37 |  | *F*(2,34) = 5.595, *p* = .008 |
| C5 (5th year) | 4 | .94 (.04) |  |
| C2 (2nd year) | 17 | .79 (.09) | C2 < C5\* |
| C3 (3rd year) | 16 | .84 (.08) |  |
|  |  |  |  |
| TRIAL-SPANISH | n = 43 |  | *t*(41) = -3.3, *p* = .002 |
| H1 (1st year) | 22 | .72 (.13) | H1 < H2\* |
| H2 (2nd year) | 21 | .85 (.13) |  |

\* indicates significant differences between respective groups, other comparisons were non-significant

Crucially, when all participants’ average proficiency scores were compared between the three trials, an ANOVA test showed no significant differences, *F*(2, 133) = 1.96, *p* = .145, = .028.

2) Full statistical information on ANCOVA results

Table 2

Trial-English (n = 55): ANCOVA results for effects of condition on learners’ comprehension, with proficiency treated as a covariate

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Df | Mean Square | F | Sig. | Partial Eta Squared |
|  |  |  |  |  |  |
| Corrected Model | 4 | .212 | 5.858 | .001 | .319 |
| Intercept | 1 | .000 | .009 | .926 | .009 |
| Proficiency | 1 | .420 | 11.615 | .001 | .189 |
| Condition | 3 | .134 | 3.715 | .17 | .182 |
| Error | 50 | .036 |  |  |  |
| Total | 55 |  |  |  |  |
| Corrected Total | 54 |  |  |  |  |

Table 3

Trial-Polish (n = 37): ANCOVA results for effects of condition on learners’ comprehension, with proficiency treated as a covariate

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Df | Mean Square | F | Sig. | Partial Eta Squared |
|  |  |  |  |  |  |
| Corrected Model | 4 | .010 | .519 | .722 | .061 |
| Intercept | 1 | .035 | 1.813 | .188 | .054 |
| Proficiency | 1 | .011 | .571 | .455 | .018 |
| Condition | 3 | .013 | .655 | .586 | .058 |
| Error | 32 | .019 |  |  |  |
| Total | 37 |  |  |  |  |
| Corrected Total | 36 |  |  |  |  |

Table 4

Trial-Spanish (n = 43): ANCOVA results for effects of condition on learners’ comprehension, with proficiency treated as a covariate

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Df | Mean Square | F | Sig. | Partial Eta Squared |
|  |  |  |  |  |  |
| Corrected Model | 4 | .190 | 6.617 | .000 | .411 |
| Intercept | 1 | 4.750E-6 | .000 | .990 | .000 |
| Proficiency | 1 | .294 | 10.223 | .003 | .212 |
| Condition | 3 | .085 | 2.952 | .045 | .189 |
| Error | 38 | .029 |  |  |  |
| Total | 43 |  |  |  |  |
| Corrected Total | 42 |  |  |  |  |

Table 5

Two-way ANCOVA results for comprehension in *–n* condition across three trials

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Df | Mean Square | F | Sig. | Partial Eta Squared |
|  |  |  |  |  |  |
| Corrected Model | 12 | .164 | 5.641 | .000 | .357 |
| Intercept | 1 | .015 | .515 | .474 | .004 |
| Proficiency | 1 | .693 | 23.811 | .000 | .163 |
| Trial | 2 | .117 | 4.030 | .020 | .062 |
| Condition | 3 | .085 | 2.952 | .045 | .189 |
| Trial \* Condition | 6 | .080 | 2.746 | .015 | .119 |
| Error | 3.552 | 122 |  |  |  |
| Total | 135 |  |  |  |  |
| Corrected Total | 134 |  |  |  |  |

3) Analysis of learners’ check marks made during listening

Following Leow et al.’s (2008) analysis in an earlier replication study based on the same design, a correlation test was performed between participants’ comprehension scores and the number of check marks, which was treated as the operationalisation of learners’ attention to form. Given that the present study involved three trials, each of them involving the measurement of comprehension via a different language, each trial was analysed separately, revealing no significant relationships (*p*s > .05). However, when the data from all our participants were considered together regardless of which language was used to measure comprehension, a weak negative correlation emerged (*rho* = -.20; *p* = .031), suggesting an inverse relationship between learners’ allocation to L2 forms and their comprehension of meaning.

Table 6

Correlations between learners’ average comprehension scores and number of check marks

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | Comprehension |  |  |
|  | Trial-English  (n = 55) | Trial-Polish   (n = 37) | Trial-Spanish  (n = 43) | All participants  (n = 135) |
|  |  |  |  |  |
| Number of check marks | ns | ns | ns | *rho* = -.20\* p = .031 |

\* indicates significant correlation