

Supplementary I

Subjective Rating of Learning and Word Learning

Additionally, we explored how learners' actual CSWL performance was related to their subjective awareness of learning by condition, as some prior studies suggest that CSWL could reflect portions of explicit processes (e.g. Li & Benitez, 2023; Poepsel & Weiss, 2014). Specifically, all adults were instructed to self-rate their confidence in word learning three times, once after each CSWL task and once at the end of the entire study, as well as to self-report their learning strategy in an open-ended question.

Method

After the training and testing phase for each CSWL condition, we also asked participants to self-report how well they had learned in the just-completed CSWL condition via a Likert scale from 0 (not learning at all) to 9 (learning a great deal), the process termed subjective rating of learning (SRL). Participants were instructed to type in the number to represent their rating on learning. Each participant subjectively rated their learning three times, once after each condition (SRL 1:1 for the 1:1 mapping condition, and SRL 2:1 for the 2:1 mapping condition) and once after completing all tasks of the study (SRL at recall). The measurement of subjective rating was added after the study was pre-registered but before data collection was conducted.

Results

Subjective rating of learning and word learning. We first asked whether learners' subjective rating of learning predicted their actual word learning in CSWL tasks by condition. Specifically, we ran two regression models in which we asked how predictors (SRL 1:1, SRL 2:1, and SRL at recall) predicted the outcomes of CSWL in the 1:1 and the 2:1 mapping condition in separate models. Table S1 depicts the regression coefficients for both models.

Results suggested that CSWL 1:1 was positively predicted by SRL 1:1 ($b = .58, p < .001$) but negatively predicted by SRL 2:1 ($b = -0.18, p = .017$). Similarly, CSWL 2:1 was positively predicted by SRL 2:1 ($b = .24, p = .004$) but not predicted by SRL 1:1 (despite the effect being marginal, $b = .16, p = .077$). SRL at recall did not serve as a significant predictor to the learning in either CSWL 1:1 ($b = -0.05, p = .514$) or CSWL 2:1 ($b = -0.13, p = .158$). In short, learners were precise in gauging how well they had learned words corresponding to conditions: the rating in the 1:1 mapping condition predicted the actual word learning in the 1:1 mapping condition (but not in the 2:1 mapping cognition), and vice versa. Yet, learners' gauging became less precise either with a time lag (at recall after completing multiple irrelevant tasks) or when asked to merge their impressions on learning from two conditions into one.

Table S1

Regression models for subjective rating of learning and CSWL

Model 1 SRL indexes predict CSWL 1:1

<i>Model</i>	R^2	ΔR^2	F	p
	.26	.24	20.22	<.001***
<i>Coefficients</i>	b	SEb	t	p
SRL 1:1	.58	.08	7.07	<.001***
SRL 2:1	-0.18	.07	-2.41	.017*
SRL at recall	-0.05	.01	-0.65	.514

Model 2 SRL indexes predict CSWL 2:1

<i>Model</i>	R^2	ΔR^2	F	p
	.08	.06	5.25	.066+
<i>Coefficients</i>	b	SEb	t	p
SRL 1:1	.16	.09	1.78	.077+
SRL 2:1	.24	.08	2.88	.004**
SRL at recall	-0.13	.09	-1.42	.158

Note. The table depicts whether and how subjective rating of learning predicted CSWL 1:1 (Model 1) and CSWL 2:1 (Model 2) for all learners. SRL 1:1 = subjective rating of word learning immediately after the 1:1 mapping

condition; SRL 2:1 = subjective rating of word learning immediately after the 2:1 mapping condition; SRL at recall = subjective rating of word learning after completing all tasks of the study. Asterisks denote significant individual coefficients and regression models. *** $p < .001$; ** $p < .01$; * $p < .05$; + $.05 < p < .10$.

Subjective rating of learning by condition and group. We second addressed whether the subjective rating of learning differed by condition and by language group. Table S2 denotes the descriptives of subjective rating of learning. Specifically, we conducted a mixed two-way ANOVA to explore how subjective rating of learning immediately after the two conditions was affected by Condition (1:1 vs. 2:1 mapping condition) and Group (monolingual vs. multilingual). As SRL at recall was shown as a non-significant predictor to both CSWL mapping conditions, we no longer considered SRL at recall in the analysis of this section. Results suggested a significant main effect of Condition such that average learners rated with a higher score of learning in the 1:1 mapping condition ($M = 4.03$, $SD = 2.87$) than that in the 2:1 mapping condition ($M = 2.99$, $SD = 2.40$), $F(1, 364) = 13.91$, $p < .001$. The main effect of Group was not significant such that monolinguals ($M = 3.43$, $SD = 2.72$) and multilinguals ($M = 3.57$, $SD = 2.68$) rated similarly across conditions, $F(1, 364) = .22$, $p = .642$. The Condition \times Group interaction was also not significant, $F(1, 364) = .00$, $p = .982$. In sum, average learners were more confident to rate regarding their learning in the 1:1 mapping condition than in the 2:1 mapping condition, and such confidence did not differ by language experience of multilingualism.

Table S2

Descriptives of subjective rating of learning

	Monolingual	Multilingual
SRL 1:1	3.95 (2.94)	4.08 (2.84)
SRL 2:1	2.92 (2.39)	3.05 (2.42)
SRL at-recall	3.81 (2.33)	3.94 (2.49)

Note. Means (standard deviations in the parenthese) are depicted for subjective rating of learning on CSWL at three different time points by group (monolingual vs. multilingual). SRL 1:1 = subjective rating on word learning immediately after the CSWL 1:1 mapping condition, on a scale from 0 (not learning at all) to 9 (learning a great deal); SRL 2:1 = subjective rating on word learning immediately after the CSWL 2:1 mapping condition.

Discussion

To conclude, the study also provides some insight into how learners' self-estimation of learning associates with their actual word learning performance. We found that learners were consistently precise in gauging how well they learned words in a condition-specific manner: the subjective rating in the 1:1 structure predicted the actual learning in the 1:1 structure but not that in the 2:1 structure, and vice versa. Consistent with the findings, prior studies suggest that learners' self-estimation can play an important role in ambiguous word learning. For instance, learners' confidence has been documented to be positively correlated with actual statistical learning performance (Li & Bentiez, 2023; Yurovsky & Yu, 2008) and to modulate the actual learning processes as learners, when being highly confident, tend to reject to change word-referent conjectures even when approved wrong (Dautriche et al., 2021). The findings also contrast some evidence which suggest that confidence may play a peripheral role for word learning, especially in the overlapping structures. For instance, learners' confidence might get boosted in overlapping structures when the overlapping pairs are differentiated by contextual cues (e.g. a pitch change, a speaker shift, or an explicit instruction, Poepsel & Weiss, 2014) or by linguistic cues (e.g. the two words for the same referent are phonotactically distinctive, Benitez et al., 2016), but their actual learning performance keeps intact despite an increase in confidence (Poepsel & Weiss, 2014; Benite et al., 2016). The current study supports that confidence and actual word learning can be *synchronous* dependent on mapping structures. Future research may

look into the robustness of the relation between confidence and word learning, and investigate under which circumstances such a relation may hold true or undergo a change.