SUPPLEMENTARY MATERIALS

to

Bordag, D. & Opitz, A. (2021). Representation and processing of overtly identical complex forms in L1 and L2: The case of conversion and its 'siblings'. *Studies in Second Language Acquisition.*

Content

I. Details of Analyses of Target Phrases: The Final Converging Model

Table S1
Reaction Times to Target Phrases, Model Statistics of the Final Model
Formula of the Final Model: logRT ~ Condition * Language + (1 | ParticipantID) + (1 | ItemID)

Random Effects	**Name**	**Variance**	***SD***
Participant	(intercept)	0.0163	0.1276
Item	(intercept)	0.0014	0.0380
Residuals		0.0459	0.2142

Number of observations: 3052, groups: Participant, 141; Item, 24

Fixed Effect	***estimate***	***SE***	**df**	***t*-ratio**	***p***
(Intercept)	6.469	0.022	289.82	300.624	<.001
Condition:identical	-0.121	0.018	2879.74	-6.567	<.001
Condition:inflected	-0.122	0.019	2880.34	-6.499	<.001
Condition:infinitive	-0.056	0.018	2879.60	-3.025	0.003
Condition:conversion	-0.061	0.019	2883.50	-3.200	0.001
Condition:noun	-0.003	0.019	2881.29	-0.179	0.858
Language:L2	0.065	0.028	326.58	2.275	0.024
Condition:identical:Language:L2	-0.035	0.026	2897.86	-1.319	0.187
Condition:inflected:Language:L2	-0.033	0.027	2891.88	-1.243	0.214
Condition:infinitive:Language:L2	-0.013	0.026	2895.88	-0.486	0.627
Condition:conversion:Language:L2	-0.024	0.028	2899.52	-0.846	0.398
Condition:noun:Language:L2	-0.005	0.027	2899.76	-0.171	0.864

Note. Linear mixed model fit by REML. t-tests use Satterthwaite's method

Fixed Effects	**sum sq**	**mean sq**	**df**	***F-ratio***	***p***	***Signif.***
Condition	9.515	1.903	5, 2896.1	41.48	<.001	***
Language	0.189	0.189	1, 139.0	4.13	.044	
Condition:Language	0.138	0.028	5, 2897.3	0.60	.699	

Note. Degrees-of-freedom method: Kenward-Roger; Signif. codes: <0.001 "***" <0.01 "**" <0.05 "*" < 0.1 "+"

II. Details of Analyses of Target Phrases: Post Hoc Analyses of L1 and L2 Participants

In the main analysis of reaction times to target phrases, no significant interaction of Language and Condition as observed ($F(5, 2897.61) = 0.60$, $p = .699$). , we additionally performed post hoc analyses for both populations separately for explorative reasons and to see whether the effect was indeed the same in L1 and L2. Analyses confirmed that Condition was significant in L1 ($F(1, 1477.83) = 19.50$, $p < .001$) and L2 ($F(1, 1403.55) = 22.07$, $p < .001$). For each subgroup of participants, pairwise comparisons of estimated means (with correction of cumulated alpha errors according to the Tukey procedure) were computed. Results are summarized in Tables S2 and S3.

Table S2

L1 Participants: Reaction Times of Target Phrases, Results of Pair-Wise Comparisons

Contrast	estimate	*SE*	df	t-ratio	p
unrelated - identical	0.121	0.017	1476.115	7.105	<.001
unrelated - inflected	0.122	0.017	1476.527	7.057	<.001
unrelated - infinitive	0.055	0.017	1476.318	3.239	0.015
unrelated - conversion	0.061	0.018	1479.854	3.450	0.008
unrelated - noun	0.002	0.017	1477.304	0.140	1.000
identical - inflected	0.001	0.017	1475.731	0.054	1.000
identical - infinitive	-0.066	0.017	1475.114	-3.897	0.001
identical - conversion	-0.060	0.017	1478.382	-3.461	0.007
identical - noun	-0.119	0.017	1476.882	-6.889	<.001
inflected - infinitive	-0.067	0.017	1476.127	-3.893	0.001
inflected - conversion	-0.061	0.018	1479.459	-3.464	0.007
inflected - noun	-0.120	0.017	1478.381	-6.839	<.001
infinitive - conversion	0.005	0.017	1478.206	0.314	1.000
infinitive - noun	-0.053	0.017	1476.640	-3.064	0.027
conversion - noun	-0.058	0.018	1480.287	-3.280	0.014

Note. p values of Tukey Contrasts with correction of cumulated alpha errors according to Tukey

Table S3

L2 Participants: Reaction Times of Target Phrases, Results of Pair-Wise Comparisons

Contrast	estimate	*SE*	df	t-ratio	p
unrelated - identical	0.157	0.020	1397.608	7.707	<.001
unrelated - inflected	0.154	0.021	1405.256	7.448	<.001
unrelated - infinitive	0.068	0.020	1403.980	3.321	0.012
unrelated - conversion	0.089	0.022	1410.604	3.962	0.001
unrelated - noun	0.006	0.021	1399.914	0.275	1.000
identical - inflected	-0.004	0.021	1399.542	-0.175	1.000
identical - infinitive	-0.090	0.020	1405.130	-4.407	<.001
identical - conversion	-0.069	0.023	1408.992	-3.046	0.029
identical - noun	-0.151	0.021	1405.213	-7.258	<.001
inflected - infinitive	-0.086	0.021	1398.666	-4.087	0.001
inflected - conversion	-0.065	0.023	1409.781	-2.842	0.052
inflected - noun	-0.148	0.021	1405.586	-6.905	<.001
infinitive - conversion	0.021	0.023	1408.571	0.928	0.939
infinitive - noun	-0.062	0.021	1405.450	-2.954	0.037
conversion - noun	-0.083	0.023	1404.354	-3.572	0.005

Note. p values of Tukey Contrasts with correction of cumulated alpha errors according to Tukey

As can be seen, the general pattern of significant differences between the six levels of "Condition" was generally identical for L1 and L2 and fully parallel with the joint analysis presented in the manuscript.

III. Analyses of Prime Phrases

In the first step, we wanted to assess the difficulty of processing of the critical forms themselves. For this purpose, we analysed the accuracy rates and reaction times of the critical forms when they appeared in the prime phrase. Since the verbs in the unrelated condition were matched with the critical verbs for both length and frequency, we included them in the analyses as well to assess the speed with which the items in this baseline condition are processed.

Accuracy

The mean accuracy rate of all primes was 92.2% (94.1% in L1, 90.2% in L2). Details are summarised in Table S4 (see also Figure S1).

Table S4

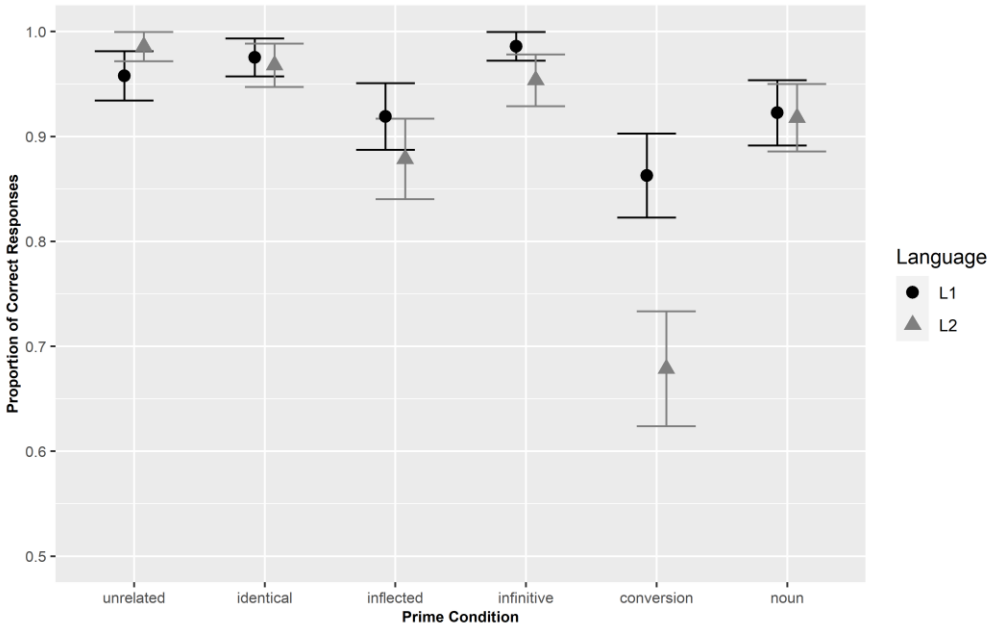
Accuracy of Responses to Prime Phrases in Percent

	unrelated	identical	inflected	infinitive	conversion	noun	*mean*
L1	95.8	97.5	91.9	98.6	86.3	92.3	*93.7*
L2	98.6	96.8	87.9	95.4	67.9	91.8	*89.7*
mean	*97.2*	*97.2*	*89.9*	*97.0*	*77.1*	*92.1*	*91.8*

Correctness of responses for primes was further analysed using a generalised linear mixed-effects model, treating the type of answer (correct vs. incorrect) as a binomial dependent variable. Prime condition and Language (L1 vs. L2) were set as fixed effects. Model comparisons yielded main effects for Prime Condition ($Chi^2(5) = 183.3$, $p < .001$) as well as a significant influence of the interaction of Prime Condition and Language ($Chi^2(5) = 23.9$, $p < .001$), but no significant main effect of Language ($Chi^2(1) = 1.70$, $p = .192$).

Figure S1

Accuracy of Judgments for Prime Phrases in L1 and L2 (Means with Standard Errors)



The significant interaction between the factors Language and Prime condition was resolved by separate analyses for both groups of participants. Post-hoc comparisons between all levels of the factor Prime condition were carried out by computing estimated means and Tukey Contrasts with correction of cumulated alpha errors (using package *emmeans* in R). The pattern of significant differences is

summarised in Tables S5 and S6. For L1 participants, the accuracy rates of the conversion condition were significantly lower than those of the unrelated, identical, and infinitive condition. Furthermore, the inflected condition led to lower accuracy rates than the infinitive condition. In L2, however, the conversion condition had significantly lower accuracy rates than all other conditions (all comparisons $p < .01$). Additionally, the inflected condition yielded lower accuracy rates than the unrelated, the identical, and the infinitive condition.

Table S5

L1 Participants: Accuracy of Prime Phrases, Results of Statistic Model

Contrast	estimate	*SE*	df	t-ratio	p
unrelated - identical	-0.562	0.487	Inf	-1.154	1.000
unrelated - inflected	0.712	0.372	Inf	1.915	0.832
unrelated - infinitive	-1.144	0.587	Inf	-1.950	0.767
unrelated - conversion	1.329	0.348	Inf	3.823	0.002
unrelated - noun	0.659	0.374	Inf	1.761	1.000
identical - inflected	1.274	0.444	Inf	2.867	0.062
identical - infinitive	-0.582	0.636	Inf	-0.916	1.000
identical - conversion	1.891	0.425	Inf	4.453	<.001
identical - noun	1.221	0.447	Inf	2.734	0.094
inflected - infinitive	-1.856	0.552	Inf	-3.362	0.012
inflected - conversion	0.617	0.284	Inf	2.172	0.448
inflected - noun	-0.053	0.317	Inf	-0.167	1.000
infinitive - conversion	2.474	0.537	Inf	4.610	<.001
infinitive - noun	1.803	0.554	Inf	3.256	0.017
conversion - noun	-0.670	0.288	Inf	-2.328	0.298

Note. p values of Tukey Contrasts with correction of cumulated alpha errors

Table S6

L2 Participants: Accuracy of Prime Phrases, Results of Statistic Model

Contrast	estimate	*SE*	df	t-ratio	p
unrelated - identical	0.839	0.621	Inf	1.350	1.000
unrelated - inflected	2.451	0.550	Inf	4.454	<.001
unrelated - infinitive	1.242	0.592	Inf	2.098	0.539
unrelated - conversion	3.885	0.539	Inf	7.213	<.001
unrelated - noun	1.917	0.564	Inf	3.398	0.010
identical - inflected	1.612	0.409	Inf	3.941	0.001
identical - infinitive	0.403	0.457	Inf	0.882	1.000
identical - conversion	3.046	0.391	Inf	7.793	<.001
identical - noun	1.078	0.419	Inf	2.575	0.150
inflected - infinitive	-1.209	0.364	Inf	-3.322	0.013
inflected - conversion	1.434	0.250	Inf	5.734	<.001
inflected - noun	-0.534	0.308	Inf	-1.735	1.000
infinitive - conversion	2.643	0.344	Inf	7.675	<.001
infinitive - noun	0.675	0.375	Inf	1.798	1.000
conversion - noun	-1.968	0.285	Inf	-6.908	<.001

Note. p values of Tukey Contrasts with correction of cumulated alpha errors.

The analyses of the accuracy rates on primes revealed that when encountering a functionally ambiguous form ending with *–en* suffix, L2 participants are inclined to identify it as a verb. When the form is presented in a syntactic context that requires a noun interpretation, participants tend to make more judgement errors. This is true especially for the conversion condition. The L2 participants judged less than 70% of the phrases comprising a converted noun as grammatically correct. L1 participants showed the lowest accuracy rates also for converted nouns (86.3%), although the difference to the other conditions was statistically less pronounced. Recall that the number of nouns and verbs was completely balanced across the experiment by means of filler phrases. A second condition that led to significantly more errors was the inflected condition. They L2 participants produced more errors in this condition than in the unrelated, identical and infinitive conditions, while for L1 participants the inflected condition statistically led to lower accuracy rates then the condition with the highest scores (i.e., the infinitive condition).

Reaction times

The analyses of reaction times of the judgment task for primes were performed only over correct responses. All statistical analyses were performed using (generalised) linear mixed-effect models employing the software R (R Core Team, 2018) with package lme4 (Bates, Mächler, Bolker & Walker, 2015). Raw data of reaction times were log-transformed. Additionally, data were winsorized with a 98% criterion for each participant, i.e. exceptionally short or long reaction times that fell below the 1st or above the 99th percentile were set to values corresponding to the 1st and 99th percentile respectively prior to further analyses. Significance of fixed-effects was determined by model comparisons using package lmerTest (Kuznetsova, Brockhoff & Christensen, 2017) using likelihood ratio tests; fixed effect terms were computed with Satterthwaite and Kenward-Roger methods for denominator degrees of freedom for F tests. All models included random intercepts for participants and items. Because no models that also included random slopes converged, models with only random intercepts were considered as the maximal random effect structure justified by the sample. Mean reaction times for prime conditions by language are given in Table S7 (see also Figure S2).

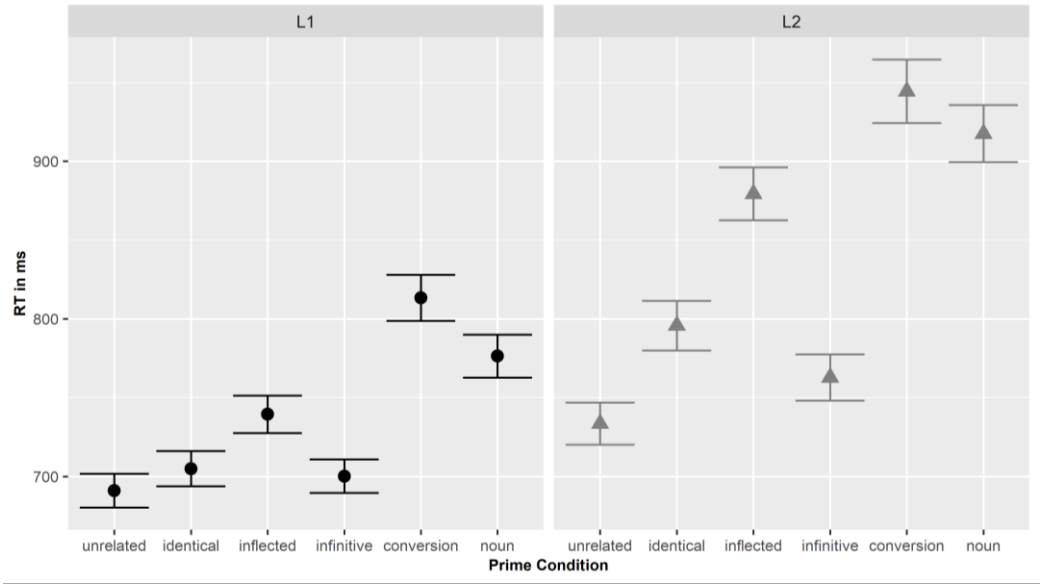
Table S7

Mean Reaction Time to Prime Phrases (in ms)

	unrelated	identical	inflected	infinitive	conversion	noun	*mean*
L1	691.1	704.1	739.0	700.6	809.0	774.0	*736.3*
L2	733.6	795.2	877.6	763.3	940.3	912.9	*837.2*
mean	*712.4*	*749.7*	*808.3*	*732.0*	*874.7*	*843.5*	*786.8*

Data were further analysed using a mixed-effects linear regression model with Language (L1 vs. L2) and Prime condition as fixed effects, and Participant and Item as random effects. Results revealed main effects for Prime condition ($F(5, 2946.6) = 65.98$, $p < .001$) and Language ($F(1, 139.3) = 24.48$, $p < .001$), and a significant interaction of the two factors ($F(5,2947.5) = 6.58$, $p < .0001$).

Figure S2
Reaction times of judgments of prime phrases in L1 and L2 (means with standard errors).



In order to resolve the interaction, differences between prime conditions were further investigated for each language group separately. For both groups of participants, the factor Condition was significant (L1: $F(5, 1500.9) = 23.41$, $p < .001$; L2: $F(5, 1433.1) = 41.43$, $p < .001$) again using pairwise comparisons of estimated means (with correction of cumulated alpha errors according to the Bonferoni procedure). The pattern of significant differences is summarised in tables S8 and S9.

Table S8
L1 Participants: Reaction Times of Prime Phrases, Results of Statistic Model

Contrast	estimate	*SE*	df	t-ratio	p
unrelated - identical	-0.007	0.007	1499.552	-0.950	1.000
unrelated - inflected	-0.028	0.007	1499.862	-3.734	0.003
unrelated - infinitive	-0.005	0.007	1499.399	-0.725	1.000
unrelated - conversion	-0.065	0.008	1503.444	-8.605	<.001
unrelated - noun	-0.044	0.007	1500.351	-5.991	<.001
identical - inflected	-0.021	0.007	1499.757	-2.810	0.075
identical - infinitive	0.002	0.007	1498.793	0.228	1.000
identical - conversion	-0.058	0.008	1502.214	-7.725	<.001
identical - noun	-0.038	0.007	1500.089	-5.079	<.001
inflected - infinitive	0.022	0.007	1499.803	3.043	0.036
inflected - conversion	-0.037	0.008	1503.061	-4.891	<.001
inflected - noun	-0.017	0.008	1501.370	-2.230	0.389
infinitive - conversion	-0.060	0.008	1501.620	-7.966	<.001
infinitive - noun	-0.039	0.007	1499.953	-5.317	<.001
conversion - noun	0.021	0.008	1504.405	2.704	0.104

Note. p values of Tukey Contrasts with correction of cumulated alpha errors according to Bonferroni

Table S9

L2 Participants: Reaction Times of Prime Phrases, Results of Statistic Model

Contrast	estimate	*SE*	df	t-ratio	p
unrelated - identical	-0.028	0.009	1431.301	-2.947	0.049
unrelated - inflected	-0.075	0.009	1431.379	-7.938	<.001
unrelated - infinitive	-0.014	0.009	1432.920	-1.541	1.000
unrelated - conversion	-0.110	0.010	1437.785	-10.751	<.001
unrelated - noun	-0.091	0.009	1432.547	-9.565	<.001
identical - inflected	-0.048	0.010	1433.020	-4.947	<.001
identical - infinitive	0.013	0.009	1430.752	1.418	1.000
identical - conversion	-0.083	0.010	1438.520	-8.006	<.001
identical - noun	-0.063	0.009	1430.429	-6.730	<.001
inflected - infinitive	0.061	0.010	1432.753	6.292	<.001
inflected - conversion	-0.035	0.011	1436.687	-3.346	0.013
inflected - noun	-0.016	0.010	1433.834	-1.610	1.000
infinitive - conversion	-0.096	0.010	1437.654	-9.250	<.001
infinitive - noun	-0.076	0.009	1431.174	-8.098	<.001
conversion - noun	0.020	0.011	1438.005	1.858	0.952

Note. p values of Tukey Contrasts with correction of cumulated alpha errors according to Bonferroni

For L1 participants, these tests revealed that there were two sub-groups of conditions. The group of slowest responses comprised the conversion and noun condition. The fastest latencies were seen for items in the unrelated, infinitive, identical and inflected conditions. The slowest responses were seen for the conversion and noun conditions. However, the inflected condition scored in-between and did not differ statistically from the fastest member of the group of slowest responses (i.e., the noun condition). For L2 participants, on the other hand, a slightly different pattern was observed. Here, the inflected condition was more distinct from the group of fastest responses (differing significantly from the unrelated, identical, and infinitive conditions). In parallel to L1, it also was not different from the noun condition, but significantly different from the conversion condition.

To sum up, the results of the analyses of both the accuracy rates and RTs of the primes revealed that participants processed with greatest ease the critical forms of 1st person plural (identical and unrelated condition) and in infinitive. The lower accuracy rate and slower responses in the 3rd person plural contexts (inflected condition) can be explained by the fact that the pronoun *sie* can mean not only "they" in German, but also "she". If interpreted as "she", the phrase, e.g., *sie spielen* ("she play") would be ungrammatical. We assume that the ambiguity of the pronoun was more difficult to resolve for the L2 participants, which caused the longer RTs and lower accuracy rates in the inflected condition (in case of wrong ambiguity resolution the phrase was indeed correctly scored as incorrect). The form most difficult to process was clearly the conversion condition and this effect was even more pronounced in L2 than in L1. The results show that in ca. 30% of the cases L2 participants did not judge the combination of the definite article *das* with the following ambiguous form as grammatical. This failure to correctly identify the grammaticality of the phrase could either be due to the fact that participants did not interpret the form as a result of conversion, i.e., a deverbal noun (but probably rather as a verb), or because they erroneously assumed a gender and/or number mismatch, for instance they could have interpreted a phrase like *das MIETEN* 'the renting' as *die MIETEN* (plural) 'the rents' that were combined with a wrong form of the article (*das* (neuter singular) instead of *die* (plural)). Even if they judged the deverbal noun in the conversion condition correctly, their response latencies were still longer than in the other conditions. Interestingly, though the responses to the concrete, countable nouns tended to somewhat lower accuracies and slower responses, at least for the accuracy rates these tendencies were not so pronounced as in the case of the conversion condition and remained only numerical. To interpret the

ambiguous form as a noun thus was not difficult per se, but only in the case when the form had the function of a deverbal noun (and especially in L2).

IV. List of Items

List of Experimental Items (with English Translations)

Verb		Countable Noun	
belegen	‘to receipt’	der Beleg	‘the receipt’
berichten	‘to report’	der Bericht	‘the report’
besuchen	‘to visit’	der Besuch	‘the visit’
beweisen	‘to prove’	der Beweis	‘the proof’
bremsen	‘to brake’	die Bremse	‘the brake’
bürsten	‘to brush’	die Bürste	‘the brush’
duschen	‘to take a shower’	die Dusche	‘the shower’
ernten	‘to harvest’	die Ernte	‘the harvest’
feiern	‘to celebrate’	die Feier	‘the celebration’
fischen	‘to fish’	der Fisch	‘the fish’
fliegen	‘to fly’	die Fliege	‘the fly’
löffeln	‘to spoon’	der Löffel	‘the spoon’
mauern	‘to mason’	die Mauer	‘the wall’
meistern	‘to master’	der Meister	‘the master’
mieten	‘to rent’	die Miete	‘the rent’
opfern	‘to sacrifice’	das Opfer	‘the victim’
pflanzen	‘to plant’	die Pflanze	‘the plant’
schrauben	‘to screw’	die Schraube	‘the screw’
schulden	‘to be in debt’	die Schuld	‘the debt’
sorgen	‘to worry’	die Sorge	‘the concern’
speisen	‘to dine’	die Speise	‘the food’
teilen	‘to divide’	der Teil	‘the part’
versuchen	‘to try’	der Versuch	‘the trial’
zelten	‘to camp’	das Zelt	‘the tent’