

Supplementary material

Stoehr, A., & Martin, C.D. (in press). Orthography affects L1 and L2 speech perception but not production in early bilinguals. *Bilingualism: Language and Cognition*.

The supplementary material includes:

Experiment I: LDT

Task instructions (Appendix S1 – page 2)

Matching of variables (Tables S1 & S2 – page 3)

Model output (Tables S3-S5 – page 4)

Experiment II: Discrimination

Model output (Tables S6 & S7 – page 5)

Experiment III: Reading-aloud

Waveform, spectrogram, and measurement interval (Figure S1 – page 6)

Matching of variables (Tables S8-S10 – page 7)

Basque-Spanish model:

Model output (Tables S11-13 – page 8)

Means and standard deviations by group, sound, and language (Table S14 – page 8)

Basque model:

Model output (Table S15 – page 9)

Means and standard deviations by group and sound (Table S16 – page 9)

Experiment I: LDT**Appendix S1: Instructions for the LDT (Experiment I)**

The original Basque instructions used in the LDT (Experiment I) with an English translation (below; adapted from Samuel & Larraza, 2015).

Original Basque version:

Ondoren, zenbait hitz entzungo dituzu: horietako batzuk euskarazko benetako hitzak izango dira eta beste batzuk, ordea, ez. Zure zeregina hitz horiek entzutea eta hitz hauek errealak diren ala ez erabakitzea izango da. Adibidez, «mahai» entzuten baduzu, egiazko hitza dela adierazi beharko duzu EZKERREKO BOTOIA sakatuz; «nahai» hitza entzuten baduzu, ostera, benetako hitza ez dela adierazi beharko duzu ESKUINEKO BOTOIA sakatuz. Adi egon, hitz faltsuek normalean benetako hitzen antza dute! Beraz, mesedez entzun arretaz eta hitz erabat zuzenak soilik onartu hitz gisara. Zure erantzuna kronometratua izango da, eta hortaz ahalik eta azkarren erantzuten saiatu behar zara. Hasieran froga bat izango duzu. Sakatu HUTSUNE-BARRA jarraitu ahal izateko.

English translation:

You are going to listen to real Basque words and to items that are not quite real Basque words. Your task is to listen to each item and decide whether it is a real Basque word or if it is not quite a real Basque word. For example, if you hear ‘mahai’ you would indicate that this was a real word by pressing the LEFT BUTTON, but if you heard ‘nahai’ you would say that this was not a real word by pressing the RIGHT BUTTON. When something is not a real word it will usually be very similar to a real word, so please listen carefully and only accept an item as a word if it is completely correct. Your response will be timed, so you need to respond as fast as possible. You will first do some practice trials. Press the SPACE BAR to continue.

Table S1. *Numerical variables matched across the three LDT stimulus lists (SD in parentheses).*

<i>Matched variables</i>	<i>List 1</i>	<i>List 2</i>	<i>List 3</i>	<i>p</i>
Position <z> ¹	1 (0)	1(0)	1(0)	1
<z> in stressed position? ²	0.64 (0.49)	0.61 (0.50)	0.68 (0.48)	.860
Syllable position in the word	1.89 (0.63)	2 (0.82)	1.89 (0.69)	.811
Number of syllables	3.11 (1.03)	3.32 (1.19)	3.25 (0.92)	.748
<z> word initial? ³	0.25 (0.44)	0.25 (0.44)	0.25 (0.44)	1
Frequency	57.66 (98.49)	63.35 (92.29)	51.54 (118.75)	.920
Orthographic length	6.89 (2.18)	7.46 (2.49)	7.14 (2.01)	.633
Number of neighbors	1.57 (2.33)	2.07 (3.11)	1.63 (3.02)	.775
AoA	3.28 (0.43)	3.28 (0.55)	3.19 (0.50)	.756
Concreteness	3.97 (0.91)	3.81 (0.65)	3.89 (0.83)	.761
OLD20 ⁴	2.45 (0.78)	2.77 (0.86)	2.74 (0.66)	.246
Spanish-Basque cognate rate	0.14 (0.12)	0.16 (0.11)	0.14 (0.13)	.745
Number of senses	2.32 (1.19)	2.44 (1.48)	2.33 (1.37)	.934
Number of Spanish translations	2.32 (0.86)	2.44 (1.48)	2.13 (1.62)	.697
Bilinguals' RT (ms)	641 (49)	648 (67)	647 (69)	.912
Bilinguals' error rate (%)	2.81 (4.28)	2.25 (3.98)	2.38 (4.30)	.876

¹ syllable onset = 1; syllable coda = 0

² stressed = 1; unstressed = 0

³ initial = 1; medial = 0

⁴ orthographic Levenshtein distance

Table S2. *Audio file durations for the three LDT conditions (SD in parentheses).*

	<i>Conditions</i>			<i>p</i>
	<i>CP</i>	<i>critical</i>	<i>control</i>	
Audio file duration (ms)	775 (160)	804 (174)	795 (178)	.540

Table S3. *Model output LDT accuracy (Experiment I). R-code: Accuracy ~ Condition * Group + (1 + Condition / Subject) + (1 + Condition + Group / Item).*

	β	SE	z	p
(Intercept)	3.589	0.259	13.839	<.001
Condition (critical vs. CP)	1.949	0.447	4.357	<.001
Condition (critical vs. control)	1.022	0.340	3.011	.003
Group	0.531	0.149	3.571	<.001
Condition (critical vs. CP)*Group	0.342	0.244	1.403	.161
Condition (critical vs. control)*Group	-0.212	0.179	1.185	>.250

Table S4. *Model output LDT reaction time (Experiment I). R-code: RT (log-transformed) ~ Condition * Group + (1 + Condition / Subject) + (1 + Condition + Group / Item).*

	β	SE	t	p
(Intercept)	7.129	0.017	427.698	<.001
Condition (critical vs. CP)	-0.030	0.008	-3.694	<.001
Condition (critical vs. control)	-0.012	0.008	-1.495	.139
Group	-0.062	0.014	-4.397	<.001
Condition (critical vs. CP)*Group	0.009	0.006	1.427	.159
Condition (critical vs. control)*Group	0.007	0.006	1.116	>.250

Table S5. *Model output LDT reaction time (Experiment I) for correct responses only. R-code: RT (log-transformed) ~ Condition * Group + (1 + Condition / Subject) + (1 + Condition + Group / Item).*

	β	SE	t	p
(Intercept)	7.129	0.018	407.381	<.001
Condition (critical vs. CP)	-0.033	0.009	-3.524	<.001
Condition (critical vs. control)	-0.015	0.008	-1.836	.070
Group	-0.064	0.015	-4.205	<.001
Condition (critical vs. CP)*Group	0.013	0.007	1.746	.087
Condition (critical vs. control)*Group	0.011	0.006	1.655	.104

Experiment II: DiscriminationTable S6. *Model output AXB task accuracy (Experiment II). R-code: Accuracy ~ Group * X-Sound * Position + (1 + X-Sound + Position | Subject).*

	β	<i>SE</i>	<i>z</i>	<i>p</i>
(Intercept)	9.289	57.704	0.161	>.250
Group	-1.946	57.702	-0.034	>.250
X-Sound	-3.972	57.699	-0.069	>.250
Position	-3.880	57.702	-0.067	>.250
Group*X-Sound	1.685	57.700	0.029	>.250
Group*Position	1.576	57.699	0.027	>.250
X-Sound*Position	2.835	57.697	0.049	>.250
Group*X-Sound*Position	-1.956	57.699	-0.034	>.250

Table S7. *Model output AXB task RT (Experiment II). R-code: RT (log-transformed) ~ Group * X-Sound * Position + (1 + X-Sound + Position | Subject).*

	β	<i>SE</i>	<i>t</i>	<i>p</i>
(Intercept)	6.004	0.049	121.800	<.001
Group	-0.056	0.049	-1.129	>.250
X-Sound	0.103	0.020	5.159	<.001
Position	-0.042	0.020	-2.111	.039
Group*X-Sound	0.038	0.020	1.896	.064
Group*Position	0.018	0.020	0.921	>.250
X-Sound*Position	0.006	0.016	0.367	>.250
Group*X-Sound*Position	0.006	0.016	0.369	>.250

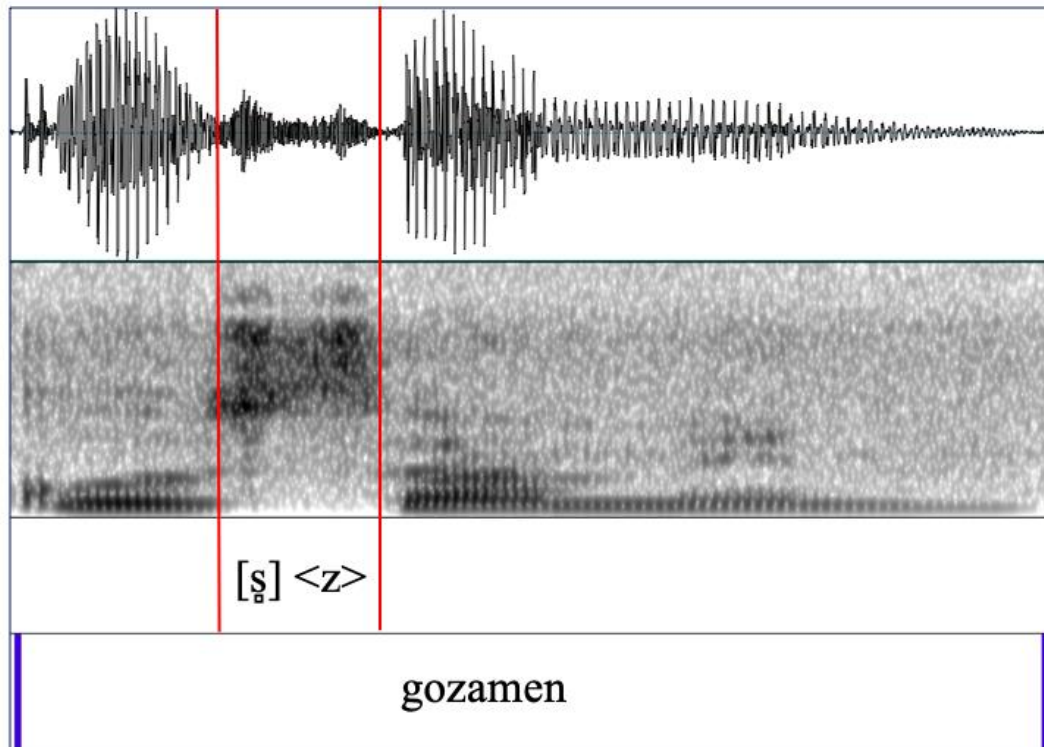
Experiment III: Reading-aloud

Figure S1. Waveform, spectrogram, and measurement interval of the fricative in the Basque word *gozamen* “enjoyment”.

Table S8. Numerical variables matched for the Basque reading-aloud task (SD in parentheses).

<i>Matched variables</i>	<i><z></i>	<i><s></i>	<i>p</i>
Frequency	22.10 (28.64)	38.42 (45.78)	.240
Orthographic length	6.45 (2.09)	6.27 (2.13)	.784
Phonological length	6.45 (2.09)	6.00 (2.11)	.507
Number of neighbors	2.65 (4.19)	1.47 (4.05)	.441
AoA	3.35 (0.50)	3.23 (0.53)	.508
Concreteness	3.86 (0.60)	4.02 (0.82)	.529
Orthographic Levenshtein distance	2.47 (0.75)	2.47 (0.72)	.977
Spanish-Basque cognate rate	0.14 (0.11)	0.21 (0.14)	.080
Number of senses	2.44 (1.36)	3.20 (2.04)	.229
Number of Spanish translations	2.13 (1.31)	3.13 (1.77)	.080
Bilinguals' RT (ms)	659 (81)	682 (73)	.424
Bilinguals' error rate (%)	5.58 (7.02)	9.05 (9.72)	.262

Table S9. Numerical variables matched for the Basque-Spanish reading-aloud tasks (SD in parentheses).

<i>Matched variables</i>	<i>Basque</i> <i><z></i>	<i>Spanish</i> <i><z></i>	<i>Basque</i> <i><s></i>	<i>Spanish</i> <i><s></i>	<i>p</i>
Orthographic length	5.92 (2.15)	6.25 (1.76)	6.08 (2.27)	6.00 (1.35)	.978
Phonological length	5.92 (2.15)	6.17 (1.75)	6.00 (2.34)	6.00 (1.35)	.991
Spanish-Basque cognate rate	0.12 (0.09)	0.17 (0.18)	0.18 (0.12)	0.20 (0.12)	.522

Table S10. Numerical variables matched for the Spanish reading-aloud task (SD in parentheses).

<i>Matched variables</i>	<i><z></i>	<i><s></i>	<i>p</i>
Spanish-Basque cognate rate	0.17 (0.18)	0.20 (0.12)	.645
Log frequency	0.70 (0.67)	1.20 (0.82)	.119
Orthographic length	6.25 (1.76)	6.00 (1.35)	.700
Number of higher frequency neighbors	0.67 (0.98)	1.67 (3.08)	.296
Orthographic Levenshtein distance	1.68 (0.37)	1.48 (0.54)	.290
Phonological length	6.08 (1.78)	6.00 (1.35)	.898
Number of syllables	2.83 (0.94)	2.67 (0.65)	.618
Position of the accented syllable	2.17 (0.94)	1.83 (0.58)	.306
Number of phonological higher frequency neighbors	1.75 (2.42)	2.50 (5.04)	.646

Table S11. *Model output reading-aloud task (Experiment III) - Basque-Spanish model. R-code: CoG ~ Sound * Group * Language + Gender + (1 + Sound + Language / Subject) + (1 + Group + Gender / Item).*

	β	SE	t	p
(Intercept)	4278.539	60.885	70.272	<.001
Sound	-266.471	3.090	-5.019	<.001
Group	-85.036	44.296	-1.920	.060
Language	-359.320	48.405	-7.423	<.001
Gender	-137.960	46.707	-2.954	.005
Sound*Group	15.108	36.699	0.412	>.250
Sound*Language	-314.314	39.847	-7.888	<.001
Group*Language	-5.607	28.981	-0.193	>.250
Sound*Group*Language	-22.973	14.719	-1.561	.121

Table S12. *Model output reading-aloud task (Experiment III) - Basque and Spanish <z>. R-code: CoG ~ Group * Language + Gender + (1 + Language / Subject) + (1 + Group + Gender / Item).*

	β	SE	t	p
(Intercept)	4027.560	78.910	51.037	<.001
Group	-63.670	59.81	-1.065	>.250
Language	-703.050	69.700	-10.086	<.001
Gender	-21.600	65.080	-0.332	>.250
Group*Language	-22.730	53.850	-0.422	>.250

Table S13. *Model output reading-aloud task (Experiment III) - Basque and Spanish <s>. R-code: CoG ~ Group * Language + Gender + (1 + Language / Subject) + (1 + Group + Gender / Item).*

	β	SE	t	p
(Intercept)	4500.870	97.670	46.083	<.001
Group	-92.220	54.830	-1.682	.097
Language	-61.520	71.390	-0.862	>.250
Gender	-201.780	55.080	-3.663	<.001
Group*Language	24.810	24.460	1.014	>.250

Table S14. *CoG in Hz by group, language, and sound (SDs in parentheses). Results are based on 12 productions per participant, language, and sound (matched on the vocalic context of the fricative).*

Language	Sound	Group	
		L1-Basque	L2-Basque
Basque	/s̺/ <z>	4,691 (1,030)	4,776 (944)
	/s̺/ <s>	4,524 (632)	4,760 (721)
Spanish	/θ/ <z>	3,294 (1,120)	3,421 (1,056)
	/s̺/ <s>	4,534 (671)	4,720 (710)

Table S15. *Model output reading-aloud task (Experiment III) - Basque <s> and <z>. R-code: CoG ~ Sound * Group + Gender + (1 + Sound | Subject) + (1 + Group + Gender | Item).*

	β	SE	t	p
(Intercept)	4656.130	65.890	70.667	<.001
Sound	65.160	54.870	1.188	.238
Group	-103.950	52.250	-1.989	.051
Gender	-181.680	52.970	-3.430	.001
Sound*Group	41.250	41.340	0.998	>.250

Table S16. *CoG in Hz by group and sound (SDs in parentheses). Results are based on 20 productions per participant and sound (matched on vocalic context of the fricative).*

Language	Sound	Group	
		L1-Basque	L2-Basque
Basque	/ʃ/ <z>	4,716 (989)	4,839 (883)
	/ʃ/ <s>	4,522 (647)	4,818 (706)