

A Guidelines for variation adaptation

A.1 Basque version

Ataza honetan testu motz batzuk hizkuntza formal/estandarretik hizkuntza informalalagoa/euskalkietara berridatzi behar dira. Bakoitzak bere hizkuntza informal/dialektalean esango lukeen bezala idaztea da helburua. Hau horrela, ondorengo aldaketak proposatzen ditugu:

- Esamolde edo hizkuntza informalagoa bilakatu.
- Ezaugarri dialektalak gehitu, bai lexiko aldetik eta bai gramatika edo fonetika aldetik.
- Hika.

Erregistroa edo dialektoak barne hartzen dituen beste edozein aldaketa ongiatorria da.
Adibidez:

Jatorrizkoa: Bi dantzari horiek dantza hunkigarria eskaini zuten herriko frontoian.

Berridatzia: Bi dantsari hoiek dantza emoziozantia eskeiñi zuten herriko frontoien.

A.2 Spanish version

En esta tarea se deben reescribir algunos textos cortos del lenguaje formal/estándar a un lenguaje más informal/dialectal. El objetivo es adaptar las frases como cada persona lo diría en su propio lenguaje dialectal. De esta manera, se propone hacer los siguientes cambios:

- A nivel de registro: Más informal, reescribiéndola de manera más coloquial
- Con rasgos dialectales, sean léxicos, gramaticales o fonéticos
- Adaptar la ortografía para que refleje vuestra pronunciación, dialecto

Cualquier otro cambio que refleje un cambio de registro o dialecto es bienvenido.

Por ejemplo:

Frase original: El amigo se quedó sin opciones cuando le dijeron que el autobús no pasaría más.

Frase adaptada: El socio se quedo botao cuando le dijeron que la guagua no pasaba ma.

A.3 English translation

This task involves rewriting short texts from formal/standard language to a more informal/dialectal language. The objective is to rewrite the sentences as each person would say them in their own dialectal language. The following changes are proposed:

- At the register level: More informal, rewriting in a more colloquial manner
- With dialectal features, whether lexical, grammatical, or phonetic
- Adapting spelling to reflect your pronunciation, dialect

Any other changes that reflect a change in register or dialect are welcome.

For example:

Original phrase: Everyone, hurry up now, dinner is about to get cold.

Adapted phrase: Y'all better hurry up now, supper's fixin' to get cold.

1140 **B Adaptation process information**

1141 **B.1 Annotator metadata**

Variable	Category	N	%
Location	Gipuzkoa	7	58.34
	Biscay	4	33.34
	Navarre	1	8.34
Age	20-30	5	41.67
	30-40	3	25.00
	40+	4	33.34
Gender	Male	5	41.67
	Female	7	58.34
Background	Linguist	8	66.67
	Non-linguist	4	33.34

Variable	Category	N	%
Location	Cuba	65	29.27
	Ecuador	50	22.52
	SpainAndalusia	55	24.77
	SpainTenerife	39	17.57
	Uruguay	13	5.85
Age	20-30		
	30-40		
	40+		
Gender	Male	2	33.34
	Female	4	66.67
Background	Linguist		
	Non-linguist		

(a) Demographic metadata of annotators. N = Count; % = Percentage

(b) Demographic metadata of annotators. N = Count; % = Percentage

Table 6: Annotator metadata

1142 **B.2 Adaptation type**

Change type	Basque		Spanish	
	N	%	N	%
re-write	18	6.04	61	27.48
dialectal	223	74.83	161	72.52
Allocutive_masc	37	12.41	-	-
allocutive_fem	20	6.71	-	-
Total	298		222	

Table 7: Number and percentage of change types in Basque and Spanish data. N: Count of examples; %: Percentage

1143 **B.3 Geographical variants distribution in data**

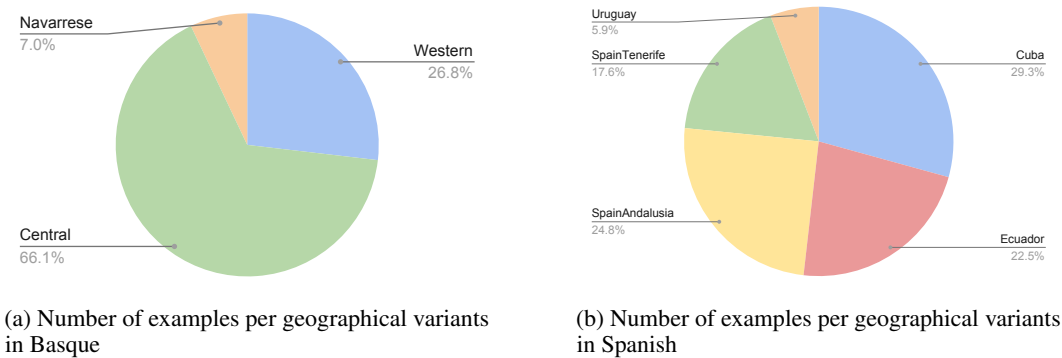


Figure 3: Geographical variant label representation in XNLIvar

Task formulation	Prompt
nli-zero	Please, answer in one word, with one of the following labels: <entailment>, <contradiction> or <neutral> Use exactly one of these three labels.
nli-few	"Say which is the inference relationship between these two sentences. Please, answer in one word, with one of the following labels: <entailment>, <contradiction> or <neutral> Use exactly one of these three labels. Here you have some examples: Postal Service were to reduce delivery frequency -> The postal service could deliver less frequently: <entailment>. This elegant spa town on the edge of the Lac du Bourget has offered cures for rheumatism and other ailments for centuries -> The town was only established in the past fifty years: <contradiction>. And while we allow people to give a kidney to their child , we do not allow them to donate their heart -> You can't always donate organs to your child: <neutral>.
qa-zero	Are these two sentences entailed, contradicted or undetermined to each other? Please, answer in one word, with one of the following labels: <entailment>, <contradiction> or <neutral> Use exactly one of these three labels.
qa-few	Are these two sentences entailed, contradicted or undetermined to each other? Please, answer in one word, with one of the following labels: <entailment>, <contradiction> or <neutral> Use exactly one of these three labels. Here you have some examples: Postal Service were to reduce delivery frequency -> The postal service could deliver less frequently: <entailment>. This elegant spa town on the edge of the Lac du Bourget has offered cures for rheumatism and other ailments for centuries -> The town was only established in the past fifty years: <contradiction>. And while we allow people to give a kidney to their child , we do not allow them to donate their heart -> You can't always donate organs to your child: <neutral>.
chain	You are an expert linguist and your task is to annotate sentences for the task of Natural Language Inference. This task consists in determining if a first sentence (premise) entails, contradicts or does not entail nor contradict the second sentence (hypothesis). Please, answer in one word, with one of the following labels: <entailment>, <contradiction> or <neutral> \n Use exactly one of these three labels \n Here you have a few examples:\n Premise: Postal Service were to reduce delivery frequency. \n Hypothesis: The postal service could deliver less frequently. \n Answer: <entailment> \n Premise: This elegant spa town on the edge of the Lac du Bourget has offered cures for rheumatism and other ailments for centuries. \n Hypothesis: The town was only established in the past fifty years. \n Answer: <contradiction> \n Premise: And while we allow people to give a kidney to their child , we do not allow them to donate their heart. \n Hypothesis: You can't always donate organs to your child. \n Answer: <neutral>

Table 8: Different task formulation prompts for generative model prompting

D Basque monolingual discriminative results

	Translate-train		
	XNLIeu	XNLIeu _{native}	XNLIeu _{var}
RoBERTa-Euscrawl	82.63	73.43	72.24
BERTeus	78.15	68.81	63.67

Table 9: Accuracy results for Basque monolingual discriminative experiments

E Accuracy results for geographical variants

	Model transfer			Translate-train			Translate-test		
	Western	Central	Navarrese	Western	Central	Navarrese	Western	Central	Navarrese
XLm-RoBERTa large	71.25	67.17	71.43	73.75	74.45	71.43	71.67	72.42	74.60
mDeBERTa	62.08	70.90	60.32	66.25	72.59	66.67	69.58	71.40	73.02

Table 10: Accuracy results for discriminative models in Basque dialects

	Model transfer				
	Cuba	Ecuador	SpainAndalusia	SpainTenerife	Uruguay
XLm-RoBERTa large	79.49	78.67	69.70	58.97	82.05
mDeBERTa	76.92	78.67	70.30	58.97	79.49
	Translate-train				
	Cuba	Ecuador	SpainAndalusia	SpainTenerife	Uruguay
XLm-RoBERTa large	79.49	84.67	70.91	61.54	82.05
mDeBERTa	73.33	81.33	69.09	53.85	87.18
	Translate-test				
	Cuba	Ecuador	SpainAndalusia	SpainTenerife	Uruguay
XLm-RoBERTa large	72.82	76.67	72.12	64.10	76.92
mDeBERTa	69.23	77.33	67.88	62.39	79.49

Table 11: Accuracy results for discriminative modes in Spanish variants