

Preliminary evaluation of EPEC-RolSem, a Basque corpus labelled at predicate level

Evaluación preliminar de EPEC-RolSem, un corpus del euskera etiquetado a nivel de predicado

Arantza Díaz de Ilarraza Sánchez

Ainara Estarrona Ibarloza

University of the Basque Country

LSI Department, UPV-EHU

Manuel Lardizabal pasealekua z/g

20018 Donostia (Gipuzkoa)

{a.diazdeilarraza}{ainara.estarrona}@ehu.es

Izaskun Aldezabal Roteta

María Jesús Aranzabe Urruzola

University of the Basque Country

Basque Philology Department, UPV-EHU

Sarriena auzoa z/g

48940 Leioa (Bizkaia)

{izaskun.aldezabal}{maxux.aranzabe}@ehu.es

Resumen: Presentamos en este artículo la evaluación preliminar de EPEC-RolSem, un corpus etiquetado a nivel de predicado con la acepción del verbo, la estructura argumental y los roles semánticos. Hemos realizado la evaluación en dos fases. Nuestra hipótesis es que con el refinamiento de criterios logrado tras la primera fase, la evaluación y los resultados de Kappa de la segunda fase mejorarán y con ello se garantizará la calidad del etiquetado posterior. Para llevar a cabo la evaluación hemos elegido 3 verbos (*adierazi*, *izan* y *etorri*) que por sus diferentes características nos permiten abarcar una amplia casuística.

Palabras clave: etiquetado de predicados, estructura argumental, roles semánticos, evaluación, Kappa, PropBank/VerbNet

Abstract: In this paper we present the preliminary evaluation of EPEC-RolSem, a corpus labelled at predicate level with verb senses, argument structure and semantic roles. We have carried out the evaluation procedure in two steps. Our hypothesis claims that with the adjustment of the criteria we get in first step, evaluation and Kappa measures will improve in second step and, thus, better quality of the tagging will be guaranteed. For this purpose, we have evaluated three verbs (*adierazi*, *izan* and *etorri*) with different properties to scope a wide casuistry.

Keywords: predicate labelling, argument structure, semantic roles, evaluation, Kappa, PropBank/VerbNet

1 Introduction and context

This paper presents the preliminary evaluation of EPEC-RolSem, a Basque corpus tagged at predicate level with verb senses, argument structure and semantic roles. It is the continuation of an ongoing work we are developing in the Ixa group within the framework of tagging corpora: EPEC corpus (*Euskararen Prozesamendurako Erreferentzia Corpora-Reference Corpus for the Processing of Basque*) (Aduriz et al., 2006) has already

been tagged morphologically and syntactically according to the dependency grammar (Basque Dependency Treebank (Aldezabal et al., 2009)), and the current aim is to incorporate predicate information to argument/adjunct candidates on the basis of the dependencies [1]. In this way, our work is in line with the general existing trends on the subject (building lexicons from tagged corpora) as shown by the corpus tagging work conducted for other languages, such as Penn Treebank (Marcus, 1994) and PropBank (Palmer et al., 2005) related to

VerbNet lexicon (Kingsbury and Palmer, 2002) and PDT, which is related to Vallex lexicon (Hajic et al., 2003). These kinds of semantic resources are essential for many computational tasks such as syntactic disambiguation and language understanding, and applications such as question answering, machine translation and text summarization.

Three basic resources are needed in a corpus annotation: the model to annotate, the guidelines to apply such a model, and the tool for tagging. We have the tool: *AbarHitz* (Díaz de Ilarraza et al., 2004), and the model was also chosen: the PropBank/VerbNet model. We conducted several analyses to find the most suitable model and we concluded that the one used by PropBank and VerbNet was suitable in the case of Basque (Agirre et al., 2006; Aldezabal et al., 2010a; Aldezabal et al., 2010c), basically for three reasons: 1) The PropBank project similar to our project is based on a syntactically annotated corpus; 2) it has been extensively used for other languages (Palmer et al., 2005; Xue, 2008; Civit et al., 2005, between others), and 3) similar criteria in order to separate senses are proposed in previous works carried out in the group; concretely in Aldezabal (2004) where a database of 100 Basque verbs (EADB–Data Base for Basque Verbs) is proposed.

Regarding the guidelines, we published the first version as an internal report based on the data obtained from the annotation of 60 verbs (Aldezabal et al., 2010b). But before continuing to tag the remaining verbs we wanted to ensure that it is complete enough, since the quality of the tagging is largely guaranteed by full tagging guidelines. For that purpose, evaluation is needed, and for the evaluation itself to be reliable we decided do it in two phases. Our hypothesis is that with the adjustment of the criteria we got in the first step, evaluation and Kappa measures will improve in the second step and, thus, better quality of the tagging will be obtained.

The paper is divided up as follows: in section 2 the structure of the tag used for the predicate labelling is explained; in section 3 the verbs worked on are described; in section 4 the evaluation procedure as well as the results and conclusions of the two phases of the evaluation are studied in depth. Finally, in section 5 we present the conclusions that are significant for our aims.

2 *The tag for predicate labeling*

We will go on to explain how we express the semantic information we assign to each syntactic dependency that is a potential verbal argument/adjunct. The semantic tag is specified as “arg_info” and comprises the following fields:

- **VN** (PropBank/VerbNet verb): the verb in English and its PropBank number, e.g.: *go_01*
- **V** (verb): dependency-relationship head, main verb
- **Element in question** (TE): argument/adjunct candidate
- **VAL** (valency): the value used to identify arguments and adjuncts: arg0, arg1, arg2, arg3, arg4, argM
- **VNrol** (VerbNet role): the VerbNet role assigned to the PropBank argument/adjunct. (Arg0: agent, experiencer...)
- **EADBrol**: the semantic role appearing in the EADB (Data Base for Basque Verbs)
- **HM** (selectional restriction): so far, only the following features are taken into consideration: [+animate], [-animate], [+human], [-human], [+concrete], [-concrete]

The *arg_info* semantic information corresponding to the word *Argentinara* (*to Argentina*) which is tagged as *nemod* can be seen in example (1):

(1) *Argentinara joan zen taldea* (The team went to Argentina)

arg_info (go_01, joan, Argentinara, Arg4, destination, end point, -)

More specific information can be found in Aldezabal et al. (2010c).

3 *Verbs worked on*

Three verbs (*adierazi* ‘to state’, *izan* ‘to be’, *etorri* ‘to come’ [2]) were selected for the evaluation work.

The verb *adierazi* (‘to state’) has only one sense, but it is very frequent in newspaper texts. As two parts of the EPEC corpus are journal texts, we can predict that the annotation of this verb will not be very complicate and that a quite significant sample

of the corpus will be easily annotated (first manually and then automatically).

The main reason for selecting the verb *izan* ('to be') was its high frequency in the corpus (15.22%). In the PropBank corpus it is not annotated because it is considered a copulative support verb; that is, it has no lexically defined semantic content and the attributes are the ones that select the arguments and their role. However, we wanted to make investigate the behaviour of such a frequent verb and to get evaluation results.

Regarding the verb *etorri* ('to come'), it is *a priori* the most difficult one to annotate. It has four senses and moreover some of the senses are not easily distinguishable. It is also used extensively in complex expressions [i.e. *bat etorri* 'to agree', *burura etorri* 'to occur to sb']. We believe that when studying a verb of this type interesting conclusions can be drawn.

Before annotating, it is necessary to ascertain the English equivalent for each sense (distinguished by numbers, i.e. *adierazi*: 1- activity) of the Basque verb. They are presented in tables 1, 2 and 3.

EADB	PropBank/VerbNet	
1- Activity	State_01	Express_01
Experiencer [3]_ERG [4] theme [-animate, -concrete]_ABS	Arg0: agent Arg1: topic Arg2: recipient Arg3: - [5] (attributive)	Arg0: agent Arg1: theme Arg2: recipient

Table1. Information for the *adierazi* verb in the EADB and in its PropBank/VerbNet equivalent.

EADB	PropBank/VerbNet
1- Location of an entity	Be_02
theme_ABS location_INE	Arg1: - (thing that is)
2- Description of an entity	Be_01
theme_ABS/ELA_KONP feature_ABS	Arg1: - (topic) Arg2: - (comment)
3- Containing of an entity [6]	Have_03
container_ABS [-animate] content_ABS [-animate]	Arg1: - (topic) Arg2: - (comment)

Table 2. Information of the *izan* verb in the EADB and in its PropBank/VerbNet equivalent.

EADB	PropBank/VerbNet	
1- Change of location	Come_01	
affected theme_ABS start point/path_ABL end point_ALA	Arg1: theme Arg2: - (extent) Arg3: - (start point) Arg4: - (end point)	
2- Creation process	Come_03	Come_09
created theme_ABS [-concrete] source_ABL [-animate] / DAT [+animate]	Arg1: theme Arg2: - (source, basis on which arg1 comes to be (not start point of motion!))	Arg1: theme Arg2: - (attribute of arg1)
3- Containing of an entity	Be_02	
content_ABS [-animate] container_INE [-animate]	Arg1: - (thing that is)	
4- Description of an entity	Be_01	
theme_ABS feature_ABS	Arg1: - (topic) Arg2: - (comment)	

Table 3. Information for the verb *etorri* in the EADB and in its PropBank/VerbNet equivalent.

4 The evaluation procedure

We carried out the evaluation in two steps. During the first step, we made an evaluation and drew some conclusions. Taking into consideration these conclusions, the guidelines have been adjusted. Then we moved on to the second evaluation and checked if the results have really improved.

In each step, 20 files for each verb have been annotated. Occurrences of the verbs vary in each file (frequency reflection, to be precise): in the first step, 27 occurrences of the verb *adierazi*, 42 of *etorri* and 74 of *izan* were found in the 60 files, and in the second one 27 occurrences of the verb *adierazi*, 42 of *etorri* and 138 of *izan*.

In the first step, the annotators independently tagged the same corpus sample.

4.1 Results of the first step

In order to calculate agreement, we first checked the sense and then the agreement when selecting the English equivalent (Table 4), because it determines the other properties (argument role, argument number, adjunct role, etc.).

VERB	VN	A1	A2	Agr.	Dis.	%
Adierazi	State_01	49	49	49	0	100
Adierazi	Express_01	5	5	5	0	100
Izan	Be_01	143	139	139	4	97.20
Izan	Be_02	12	14	12	2	85.71
Izan	Have_03	27	29	27	2	93.10
Etorri	Come_01	29	26	22	11	89.65
Etorri	Come_03	2	7	0	9	0
Etorri	Be_01	2	0	0	2	0

Table 4. Selected senses and degree of agreement between annotators.

We measured the same with Cohen's Kappa (Carletta, 1996). Table 5 shows the results.

adierazi	1.000
izan	0.939
etorri	-0.120

Table 5. Cohen's Kappa on selected senses.

Tables 4 and 5 show that altogether, there is considerable inter-annotator agreement when selecting the sense, and consequently, the English equivalent. Yet for the verb *adierazi*, there is not a single disagreement. As the verb only has one sense, we had anticipated it. Even though two English equivalents are used to translate that sense, there was no problem for selecting one or the other.

In the case of the verb *izan*, the circumstances were slightly more difficult because it has three senses and is a copulative verb, yet it can be seen that the level of agreement is fairly high, which was surprising as we had expected more difficulties regarding this verb. Even if it is a copulative verb, the senses are obviously clearly distinguished in the sentences. The same equivalent was selected 178 times, and a different one 4 times. One of the cases in which a different equivalent was selected by the annotators is illustrated in example (2):

(2) *Kasparovi kendu dio Kramnik gazteak koroa, hamabost urtean harena izan ostean.*

Lit. The young Kramnik took the crown to Kasparov, after being his during fifteen years

One annotator selected the equivalent "have_03" and the other "be_01": while one annotator considered the verb to denote the possession of an object (be_01), the other considered it to denote containing (have_03).

Yet in the case of *etorri* the same sense or equivalent was selected 22 times, and on 11 occasions they did not agree when selecting the equivalent; consequently, Kappa is also very low. Moreover, it has to be mentioned that the agreement cases regard the first sense; in the other two senses (that appeared in the text) there is no agreement. That fact suggests to us that the limit of these two senses is not clear enough.

We can see a case of disagreement in example (3):

(3) *Oso gaztetatik datorchio xakeko zaletasuna.*

Lit. Since he was very young comes the chess fondness.

He got fond of chess at a very young age.

In this case one annotator chose the first sense (change of location) and then selected "come_01", and the other one chose the second sense (creation process) and then selected "come_03". Since *tendency* is an abstract noun it turns out to be much more difficult to assign one or the other sense.

In addition, we obtained other data with Cohen's Kappa: the agreement for verb sense and valency (Table 6), and the agreement for verb sense, valency and semantic role (Table 7).

Tables 6 and 7 show that when taking into account the semantic role the Kappa of *adierazi* and *izan* decrease slightly. That is quite logical since we took into account more variables. However, we checked the results manually, and we were able to see that the disagreements occur when assigning the role to the adjuncts.

English equivalent + valence	
Adierazi	1.000
Izan	0.950
Etorri	0.232

Table 6. Kappa measures taking into account two variables: the English equivalent and the valence.

English equivalent + valency + role	
Adierazi	0.783
Izan	0.846
Etorri	0.231

Table 7. Kappa measures taking into account three variables: the English equivalent, the valency and the semantic role.

In the next section we will explain the conclusions we reached during this first step.

4.2 Conclusions of the first step

Regarding the coverage of the guidelines it can be said that there is a gap in the modifier section, which will require refining the criteria; but it also has to be said that some disagreements are unavoidable because in some cases modifiers are naturally ambiguous. For instance: in *hitzaldian adierazi* (*express in the speech*), does the INE (inessive) express time or place? Or in *amaitzear zegoela* (*being about to finish*): Should it be understood as manner or time? Depending on the annotator's understanding, both are valid. Therefore, disagreements of this nature did not strike us as significant, bearing in mind that there will always be a percentage of disagreement.

Multi-lexical units (MLU) were another source of disagreements. Although some MLUs are recognized in previous phases of the tagging, there are still gaps in this area. For instance, in the example *Sharonen jarrera probokatzailea zertara datorren galdetu zuen Mubarakek* (Lit. Mubarak asked what does the Sharon's provocative attitude come for), one annotator considered *zertara etorri* ('what come for') as MLU and the other did not.

Another problem are the vocatives. The guidelines do not specify whether vocatives need to be tagged, which is why in the only case a vocative occurred one annotator tagged it and the other did not. It is true that there are very few occurrences of the vocative in the corpus, but in our view it needs to be specified in the guidelines.

On the other hand, although the annotators agree when selecting the English equivalent the disagreements appear when tagging other features, such as the number of the argument and the role. In some cases, one annotator followed EADB and the other one PropBank. Moreover, there were confusions when applying the guidelines' criteria (both from EADB and PropBank). We observed it above all in the verb *etorri*.

For instance, in PropBank "come_01" has an "extent" Arg2, which is not possible in Basque. Although the role does not exist for this verb, one annotator continued using the numbered Arg2 for another existing role

(Arg2: start point), while the other annotator left aside also the numbered argument, maintaining the argument-role link of PropBank (Arg3: start point). It has to be pointed out that the "extent" argument has the "rare" mark in PropBank, which shows that it is not a common argument in English either.

Other disagreement occurs when tagging Arg1. PropBank always assigns the role "theme" to Arg1 but we did not apply this criterion in our guidelines, that very closely follow the PropBank guidelines (Babko-Malaya, 2005). As far as we understand, the argument level and the role level are independent one from each other. For an unaccusative verb like "come_01", where only the intransitive variant is possible, we considered that the entity that does and undergoes the actions is the same; thus, we propose to tag it as Arg0, unlike the causative/inchoative verbs like *break*, where the "theme" is always Arg1 even the "cause" (Arg0) is not explicit in the sentence. In these cases, one annotator did not follow our guidelines and tagged it as in PropBank.

The main conclusion that arises is the importance of fully editing the verb entry before starting annotating: not only the English equivalent for the sense must be clear, but also the numbered arguments and the assignment of the roles. And this is exactly what we did in the second step.

On the other hand, our main goal was to prove the appropriateness of the guidelines, and after analysing the results, we detected some gaps. We need to:

- define the roles for the adjuncts more clearly,
- clarify what to do with vocatives and
- adjust the criteria for the adaptation of the PropBank/VerbNet model.

Finally, we have to mention the problematic issue of the MLUs. It is clear that it is a slippery field, and it is difficult to propose cues to guide the detection. The annotators should take into account that not all possible MLUs are previously detected and that they will be probably the ones that will detect the new ones.

4.3 Results and conclusions of the second step

As we pointed out, the first task of the second step is the editing process. An entry for each sense of the Basque verb was prepared in PropBank/VerbNet style by the two annotators. For this purpose the annotators first have to properly understand the sense of each verb; second, they need to have a clear idea about the argument structure; and third, they have to decide on an acceptable English translation.

We show the proposed entries in the following tables (Tables 8, 9 and 10).

V: adierazi
 VN: state_01 / express_01
 VAL: Arg0, VNrol: agent, EADBrol: experiencer_ERG
 VAL: Arg1, VNrol: topic, EADBrol: gaia_ABS/ELA_KONP, HM: -animate, -concrete

Table 8. The entry for the verb *adierazi* in PropBank/VerbNet style.

1- Location of an entity
 V: izan
 VN: be_02
 VAL: Arg1, VNrol: theme, EADBrol: theme_ABS
 VAL: Arg2, VNrol: location, EADBrol: location_INE

2- Description of an entity
 V: izan
 VN: be_01
 VAL: Arg1, VNrol: topic, EADBrol: theme_ABS/ELA_KONP
 VAL: Arg2, VNrol: attribute, EADBrol: feature_ABS

3- Possession of an entity
 V: izan
 VN: have_03
 VAL: Arg0, VNrol: theme, EADBrol: container_ERG
 VAL: Arg1, VNrol: theme, EADBrol: content_ABS

Table 9. The entry for the verb *izan* in PropBank/VerbNet style.

1- Change of location
 V: etorri
 VN: come_01
 VAL: Arg0, VNrol: theme, EADBrol: affected theme_ABS
 VAL: Arg1, VNrol: source/path, EADBrol: start point/path_ABL
 VAL: Arg2, VNrol: destination, EADBrol: end point_ALA

2- Creation process
 V: etorri
 VN: come_03 / come_09 (*come out*)
 VAL: Arg0, VNrol: theme, EADBrol: created theme_ABS, HM: -concrete
 VAL: Arg1, VNrol: location, EADBrol: source_ABL, HM: -animate/_DAT, HM: +animate

3- Containing of an entity
 V: etorri
 VN: be_02
 VAL: Arg0, VNrol: theme, EADBrol: content_ABS, HM: -animate
 VAL: Arg1, VNrol: location, EADBrol: container_INE, HM: -animate

4- Description of an entity
 V: etorri
 VN: be_01
 VAL: Arg0, VNrol: topic, EADBrol: theme_ABS
 VAL: Arg1, VNrol: attribute, EADBrol: feature_ABS

Table 10. The entry for the verb *etorri* in PropBank/VerbNet style.

Afterwards, each annotator tagged the same sample of the corpus without commenting to each other on anything.

As in the first step, we checked firstly the sense and consequently the agreement existing when selecting the English equivalent (Table 11):

VERB	VN	A1	A2	Agr.	Dis.	%
Adierazi	State_01	45	41	41	4	65.79
Adierazi	Express_01	18	22	18	4	81.81
Izan	Be_01	220	228	219	10	96.05
Izan	Be_02	35	26	26	9	74.28
Izan	Have_03	36	37	36	1	97.29
Etorri	Come_01	28	32	28	4	87.5
Etorri	Come_03	8	8	8	0	100
Etorri	Be_01	2	2	2	0	100

Table 11: Selected senses and degree of agreement between annotators.

Confirming our hypothesis, the results of the agreement in the case of *etorri* have evidently improved.

Cohen's Kappa (Carletta, 1996) measures also show the same (Table 12).

adierazi	0.854
izan	0.910
etorri	0.781

Table 12. Cohen's Kappa on selected senses.

There is a remarkable difference for the verb *etorri* between the first step (-0.120) and the second one (0.781) when annotation the corpus after carrying out the editing task.

On the other hand, there was a small decrease in the results of *adierazi*, since in some cases one annotator selected the “state_01” equivalent and the other annotator the “express_01” one (see Table 11). However, we do not think it is such an important fact because really they agreed in choosing the sense and consequently there is no difference in argument structure neither in the roles.

When taking into account also the valency, the agreement for *adierazi* and *izan* decreases slightly (Table 13). In the case of *adierazi* the reason is the same: differences in assigning the English equivalent. In the case of *izan*, the difference with respect to the first step is not significant. In the case of *etorri*, the improvements are again confirmed.

Adding semantic role information in the variables, all the results with respect to the first step improve for the three verbs (Table 14). In the first step we saw that when taking into account also the role the results deteriorate because of differences in assigning the role to the adjuncts. Refining the criteria for the role assignment of the adjuncts in the guidelines seems to be effective.

English equivalent + valence	
adierazi	0.922
izan	0.930
etorri	0.818

Table 13. Kappa measures taking into account two variables: the English equivalent and the valence.

English equivalent + valency + role	
adierazi	0.808
izan	0.869
etorri	0.704

Table 14. Kappa measures taking into account three variables: the English equivalent, the valency and the role.

The main conclusion we obtained from this second step is a confirmation of the importance of the full edition task before carrying out the annotation.

5 General conclusions

After the improvements of the first evaluation phase we achieved a good level of agreement. As a conclusion, first, we can confirm that the PropBank/VerbNet model

has been found to serve our purposes, although we have to make several adaptations to the model, and second, after including the improvements of the first evaluation (better definition of adjunct’s role assignment, resolution of the vocatives and adjustment of the criteria when applying the PropBank/VerbNet model) the guidelines have an adequate coverage and quality.

However, due to the limits of the guidelines, the need of a verb by verb edition task has become apparent. At the most, the guidelines can provide with the most general criteria possible. This is, in fact, the most important conclusion. In the future, we plan to make another evaluation when we have annotated a larger sample of the corpus.

Finally, we have seen that this semantic tagging will enable us to improve work previously carried out. We have seen that in some cases the dependencies are not correct or that the MLUs do not come identified, so we have an opportunity to resolve these errors in the current phase.

Acknowledgements

This research is supported by the University of the Basque Country (GIU09/19), the Basque Government (IT344-10) and The Ministry of Science and Innovation of the Spanish Government (EPEC-Rolsem project (FFI2008-02805-E/FILO) (Complementary Action) and AncoraNet project (FFI2009-06497-E)). We would especially like to thank Oier Lopez de la Calle for helping us to apply Cohen's Kappa and write the final version of the paper.

Notes

- [1] At semantic level, so far, the nouns were tagged by means of Euskal WordNet senses (Pociello et al., 2010).
- [2] These translations are for guidance.
- [3] These roles exist in Basque, in fact, but we translated them for guidance.
- [4] In Basque we also determine the declension case in which the role realizes. The meanings of the abbreviations that appear in the examples are the following: ERG (ergative), ABS (absolutive), ELA_KONP (-ela completive), ALA (allative), ABL (ablative), INE (inessive). When there is a syntactic alternation in

- the arguments (i.e. between ABS and ELA_KONP), we expressed it by means of “/”.
- [5] When VerbNet does not provide any role for the argument we mark it with “-“ and the role that PropBank proposes in parenthesis.
- [6] It has to be pointed out that no example of this sense appeared in the files selected for evaluation purposes.

References

- Aduriz, I., M. J. Aranzabe, J. M. Arriola, A. Atutxa, A. Díaz de Ilarraza, N. Ezeiza, K. Gojenola, M. Oronoz, A. Soroa and R. Urizar. 2006. Methodology and steps towards the construction of EPEC, a corpus of written Basque tagged at morphological and syntactic levels for automatic processing. In Andrew Wilson, Paul Rayson and Dawn Archer (eds.), *Corpus Linguistics Around the World*. Book series: Language and Computers. Vol. 56, 1-15. Rodopi (Netherlands).
- Agirre, E., I. Aldezabal, J. Etxeberria and E. Pociello. 2006. A Preliminary Study for Building the Basque PropBank. *Proceedings of the 5th International Conference on Language Resources and Evaluations (LREC)*. Genoa, Italy.
- Aldezabal, I. 2004. *Aditz-azpikategorizazioaren azterketa. 100 aditzen azterketa zehatza, Levin (1993) oinarri harturik eta metodo automatikoak baliatuz*. Leioa (Bilbao), University of Basque Country thesis.
- Aldezabal, I., M. J. Aranzabe, J. M. Arriola and A. Díaz de Ilarraza. 2009. Syntactic annotation in the Reference Corpus for the Processing of Basque (EPEC): Theoretical and practical issues. *Corpus Linguistics and Linguistic Theory* 5-2, 241-269. Mouton de Gruyter. Berlin-New York.
- Aldezabal, I., M. J. Aranzabe, A. Díaz de Ilarraza and A. Estarrona. 2010a. Building the Basque PropBank. In Nicoletta Calzolari, Khalid Choukri, Bente Maegaard, Joseph Mariani, Jan Odjik, Stelios Piperidis, Mike Rosner and Daniel Tapias (eds.), *Proceedings of the Seventh Conference on International Language Resources and Evaluation (LREC 2010)*, 1414-1417, European Language Resources Association (ELRA), Valletta (Malta).
- Aldezabal, I., M. J. Aranzabe, A. Díaz de Ilarraza, A. Estarrona, K. Fernández and L. Uria. 2010b. EPEC-RS: EPEC (Euskararen Prozesamendurako Erreferentzia Corpusa) rol semantikoekin etiketatze eskuliburua [Guidelines to tag semantic roles in the EPEC corpus (the Reference Corpus for the Processing of Basque)]. Internal Report, UPV / EHU / LSI / TR 02-2010.
- Aldezabal, I., M. J. Aranzabe, A. Díaz de Ilarraza, A. Estarrona and L. Uria. 2010c. EusPropBank: Integrating Semantic Information in the Basque Dependency Treebank. In Alexander Gelbukh (ed.), *Lecture Notes in Computer Science (LNCS) n° 6008, Computational Linguistics and Intelligent Text Processing*, 60-73, Springer, Berlin-Heidelberg-New York.
- Babko-Malaya, O. 2005. PropBank Annotation Guidelines. In <http://verbs.colorado.edu/~mpalmer/projects/ace/PBguidelines.pdf>
- Carletta, J. 1996. Assessing agreement on classification tasks: The kappa statistic. *Computational Linguistics*, 22(2): 249–254.
- Civit, M., I. Aldezabal, E. Pociello, M. Taulé, J. Aparicio and L. Màrquez. 2005. 3LBLEX: léxico verbal con frames sintáctico-semánticos. In *XXI Congreso de la SEPLN*. Granada, Spain
- Díaz de Ilarraza, A., A. Garmendia and M. Oronoz. 2004. Abar-Hitz An Annotation Tool for the Basque Dependency Treebank, *Language Resources and Evaluation Conference (LREC 2004)*, Lisbon, Portugal.
- Hajic, J., J. Panevová, Z. Urešová, A. Bémová, V. Kolárová and P. Pajas. 2003. PDT-VALLEX: Creating a Largecoverage Valency Lexicon for

- Treebank Annotation. In *Proceedings of the Second Workshop on Treebanks and Linguistic Theories*, 57–68. Sweden.
- Kingsbury, P. and M. Palmer. 2002. From Treebank to PropBank. In *Proceedings of the 3rd International Conference on Language Resources and Evaluation (LREC-2002)*. Las Palmas, Spain.
- Marcus, M. 1994. The Penn TreeBank: A revised corpus design for extracting predicate argument structure. In *Proceedings of the ARPA Human Language Technology Workshop*. Princeton, NJ.
- Palmer, M., D. Gildea and P. Kingsbury. 2005. The Proposition Bank: A Corpus Annotated with Semantic Roles. In *Computational Linguistics Journal*, 31:1.
- Pociello, E., E. Agirre and I. Aldezabal. 2010. Methodology and Construction of the Basque WordNet. *Language Resources and Evaluation (LRE) Journal*.
- Xue, N. 2008. Labeling Chinese predicates with semantic roles. *Computational Linguistics*, 34(2): 225-255