

Coreference Resolution for the Basque Language with BART

**Ander Soraluze, Olatz Arregi,
Xabier Arregi, Arantza Díaz de Ilarraza**
University of the Basque Country
Donostia - San Sebastián, Spain
{`ander.soraluze, olatz.arregi`
`xabier.arregi, a.diazdeilarraza`} @ehu.eus

Mijail Kabadjov, Massimo Poesio
University of Essex
Colchester, UK
{`malexa, poesio`}
@essex.ac.uk

Abstract

In this paper we present our work on Coreference Resolution in Basque, a unique language which poses interesting challenges for the problem of coreference. We explain how we extend the coreference resolution toolkit, BART, in order to enable it to process Basque. Then we run four different experiments showing both a significant improvement by extending a baseline feature set and the effect of calculating performance of hand-parsed mentions vs. automatically parsed mentions. Finally, we discuss some key characteristics of Basque which make it particularly challenging for coreference and draw a road map for future work.

1 Introduction

Basque is a language spoken by nearly three quarters of a million people, most of which live in the Basque country, a region spanning parts of northern Spain and southwestern France. One of the most surprising findings about the Basque language is that it cannot be linked with any of its Indo-European neighbours in Europe and, hence, has been classified as a language isolate. It differs considerably in grammar from the languages spoken in surrounding regions. It is an agglutinative, head-final, pro-drop, free-word order language (Laka, 1996).

Naturally, the Basque language has also inspired a lot of work in Computational Linguistics with tools for automatically processing it becoming increasingly available (Alegria et al., 1996; Alegria et al., 2002; Alegria et al., 2003; Aduriz and Díaz de Ilarraza, 2003; Alegria et al., 2008). However, as it is

the case with most less-resourced languages, there are tools for the core processing levels, such as tokenisation, sentence splitting, morphological analysis, syntactic parsing/chunking, but much less so for higher semantic levels required in end goal applications such as Question Answering (Morton, 2000), Text Summarisation (Steinberger et al., 2007) or Information Extraction (Def, 1995; Hirschman, 1998). One such intermediate problem which has been underresearched for Basque, and hence, no readily usable tools are publicly available yet, is that of Coreference Resolution (Poesio et al., 2016).

However, preliminary work on Coreference for Basque is starting to emerge (Soraluze et al., 2015), and in this paper we describe our work on extending the coreference resolution toolkit, BART¹ (Versley et al., 2008) to the Basque language. BART benefits from an open architecture and provides a mechanism through language plugins which makes it particularly suitable for adaptations to new languages, and it attained good performance in the shared task on Multilingual Coreference at CoNLL 2012 (Uryupina et al., 2012).

For our experiments we use the EPEC corpus annotated for coreference (Aduriz et al., 2006) and we run experiments across two dimensions. First, we use a baseline model based on (Soon et al., 2001) vs. a model that includes extra features reliably extracted for Basque with the tools at hand. Second, we measure performance on hand-parsed mentions vs. performance on automatically parsed mentions which illustrates the effect of pre-processing quality on the end results.

¹<http://www.bart-coref.eu/>

One of the key challenges that the Basque language introduces for Coreference is that it uses a genderless system for pronouns. In our experiments we look in more depth around this issue and show the challenges it presents as well as suggest viable solutions to model it with machine learning techniques.

The remainder of this paper is organised as follows: Section §2 briefly surveys related work, Section §3 gives details of EPEC, a coreference corpus, Section §4 describes the extension of BART to Basque, Section §5 presents results and provides a discussion on the challenges for coreference in Basque, and towards the end we draw conclusions and pointers to future work.

2 Related Work

Preliminary work on Coreference for Basque was done by (Soraluze et al., 2015) where they adapt the Stanford coreference resolution system (Lee et al., 2013) to Basque. And there has been a lot of work on extending the BART coreference toolkit to languages other than English. (Poesio et al., 2010) extend it to Italian using the Evalita corpus of Wikipedia articles (Broscheit et al., 2010) work on German using the TüBa-D/Z coreference corpus, (Kopeć and Ogrodniczuk, 2012) develop the Polish plug-in using a subset of the National Corpus of Polish, and finally (Uryupina et al., 2012) run experiments on Arabic and Chinese.

3 Annotated Corpus of Basque

EPEC (Reference Corpus for the Processing of Basque) (Aduriz et al., 2006) is a 300,000 word sample collection of standard written Basque that has been manually annotated at different levels (morphology, surface syntax, phrases, etc.). The corpus is composed by news published in *Euskaldunon Egunkaria*, a Basque language newspaper. It is aimed to be a reference corpus for the development and improvement of several NLP tools for Basque.

Recently, mentions and coreference chains were also annotated by two linguists in a subset of the EPEC corpus which is composed of about 45,000 words. First, automatically annotated mentions obtained by our mention detector were corrected; then, coreferent mentions were linked in clusters. The

mention detector is a set of hand-crafted rules that have been compiled into Finite State Transducers (FST). The FSTs match chunks and clauses provided by the preprocessing tools and identify the mentions and their boundaries. Further discussion about the FSTs’ behaviour can be found in (Soraluze et al., 2012).

All the annotation process has been carried out using the MMAX2 annotation tool (Müller and Strube, 2006). The coreference annotation of the EPEC corpus is explained more in detail in (Ceberio et al., 2016).

To adapt BART to Basque, we divided the dataset into three main parts: one for training the system, the other for tuning, and the last for testing. More detailed information about the three parts can be found in Table 1.

	Words	Mentions	Clusters	Singletons
Train	23520	6525	1011	3401
Devel	6914	1907	302	982
Test	15949	4360	621	2445

Table 1: EPEC-coref corpus division information.

4 Extending BART to Basque

BART was originally created for English, but its flexible modular architecture ensures its portability to other languages.

BART consists of five main components: preprocessing pipeline, mention factory, feature extraction module, decoder and encoder. Furthermore, an additional independent *Language Plugin* module handles language specific information and is accessible from any component.

In the adaptation process of BART, we used a preprocessing pipeline of Basque linguistic processors, developed the *Basque Language Plugin* and added new features for coreference resolution specifically geared towards Basque.

4.1 Preprocessing and Mention Detection

The preprocessing pipeline takes raw texts and applies a series of Basque linguistic processors to analyse the texts: i) A morphological analyser that performs word segmentation and PoS tagging (Alegria et al., 1996), ii) A lemmatiser that resolves the ambiguity caused at the previous phase (Alegria et al.,

2002), iii) A multi-word item identifier that determines which groups of two or more words are to be considered multi-word expressions (Alegria et al., 2004), iv) A named-entity recogniser that identifies and classifies named entities (person, organisation, location) in the text (Alegria et al., 2003), v) A chunker, an analyser that identifies verbal and nominal chunks based on rule-based grammars (Aduriz and Díaz de Ilarraza, 2003), vi) A clause tagger, that is, an analyser that identifies clauses, combining rule-based-grammars and machine learning techniques (Alegria et al., 2008).

After the preprocessing step, mentions that are potential candidates to be part of coreference chains are identified using the mention detector explained in Section 3.

Finally, the linguistic information obtained by the preprocessing tools and the mentions identified by the mentions detector are stored in stand-off format of the MMAX2 annotation tool (Müller and Strube, 2006) that BART uses.

4.2 Basque Language Plugin

Developing a Basque language plugin for BART involved building on the system’s already existing language plugins, and then translating closed-class words such as pronouns, mapping key part-of-speech tags and adapting lower-level heuristics for finding the head noun in noun phrases, person and number identification, as well as reading features made available by the preprocessing tools.

4.3 Feature engineering for Basque

Some kind of linguistic information from the mention is used by all the features implemented in BART. MentionFactory computes these properties when a language is supported by BART. In the case of a new language, such as Basque, they should be provided as part of the mention representation computed by external preprocessing facilities. So, we added in the MMAX2 files relevant features for coreference resolution in Basque, as are number and lemma.

For our experiments, we trained BART with two different models. The first one, is a simple model, presented by (Soon et al., 2001).² The second one,

²Due to the way we integrated the preprocessing pipeline for

is an improved version of the first one where more Basque oriented features have been added. The features used in each model are presented in Table 2.

In the two models, gender agreement does not cause any improvement in the scores, as Basque is genderless.³

At this point the proposed new features to handle the specificity of Basque are not new and have also been used for other languages (see (Poesio et al., 2016) for details).

5 Experimental Results

We have tested the two models presented in Subsection 4.3 in two different environments. In the first one automatically detected mentions are provided to the models and in the second one the mentions are gold.⁴

The metrics used in our evaluations are MUC (Vilain et al., 1995), B^3 (Bagga and Baldwin, 1998), $CEAF_e$ (Luo, 2005), $CEAF_m$ (Luo, 2005), and BLANC (Recasens and Hovy, 2011). The scores have been calculated using the reference implementation of the CoNLL scorer (Pradhan et al., 2014).

Table 3 presents the results obtained by the two models when automatic mentions are used.

		R	P	F_1
Mention Detection		72.91	74.69	73.79
MUC	Soon	18.37	67.23	28.86
	Basque	35.44	45.53	39.86
B^3	Soon	53.96	72.85	62.00
	Basque	58.10	65.27	61.48
$CEAF_m$	Soon	57.50	58.90	58.19
	Basque	58.67	60.10	59.38
$CEAF_e$	Soon	67.42	52.93	59.31
	Basque	61.63	58.15	59.84
BLANC	Soon	32.29	62.47	36.46
	Basque	38.70	48.81	42.41
CONLL	Soon	-	-	50.05
	Basque	-	-	53.72

Table 3: Scores with automatic mentions.

In the case of automatically detected mentions, Basque model outperforms the Soon baseline model

Basque with BART, at this stage we were unable to incorporate all features in the original (Soon et al., 2001) model.

³We maintain this feature with the aim of not modifying the (Soon et al., 2001) model.

⁴Since the official CoNLL scorer is used for the evaluation, it also takes care of the alignment between automatically detected mentions and gold ones.

Features		Baseline	Basque
Gender	M_i and M_j agree in gender	✓	✓
Number	M_i and M_j agree in number	✓	✓
Alias	Matches abbreviations and name variations	✓	✓
StringMatch	M_i and M_j have the same surface form	✓	✓
SemClassAgree	Assesses the semantic compatibility of M_i and M_j	✓	✓
Appositive	M_i and M_j are in apposition structure	✓	✓
DistanceSentence	Distance in sentences between M_i and M_j	✓	✓
LemmaMatch	M_i and M_j have the same surface lemma	×	✓
HeadMatch	M_i and M_j have the same head	×	✓
StringKernel	Computes the similarity M_i and M_j strings	×	✓
DistanceMarkable	Distance in markables between M_i and M_j	×	✓
HeadPartofSpeech	M_i and M_j head PoS are the same	×	✓

Table 2: Features used for Coreference Resolution in our experiments. M_i is a candidate antecedent and M_j is a candidate anaphor.

according to F_1 on all the metrics except B^3 . In CoNLL metric, Basque model has a score of 53.72, which is 3.67 points higher than Soon Baseline, which scores 50.05.⁵

Scores obtained when gold mentions are provided are shown in Table 4.

Mention Detection		R	P	F_1
		100	100	100
MUC	Soon	23.62	78.66	36.34
	Basque	49.49	57.28	53.10
B^3	Soon	74.66	98.00	84.75
	Basque	81.21	87.78	84.37
$CEAF_m$	Soon	75.58	75.58	75.58
	Basque	76.59	76.59	76.59
$CEAF_e$	Soon	91.11	70.29	79.35
	Basque	82.10	77.64	79.81
BLANC	Soon	57.08	89.79	61.68
	Basque	66.78	75.99	70.34
CoNLL	Soon	-	-	66.81
	Basque	-	-	72.42

Table 4: Scores with gold mentions.

When gold mentions are used the Basque model also outperforms the Soon baseline according to all the metrics, except B^3 . The official CoNLL metric is outperformed by 5.61 points.

Comparing the results obtained when gold mentions are used with those obtained with the automatic mentions, there is a considerable difference. CoNLL F_1 of Soon baseline is 50.05 when automatic mentions are provided, while providing gold mentions this value raises to 66.81, an increase of 16.76. Similar increase in CoNLL F_1 happens with the Basque model. In this case, there is an increase

⁵The CoNLL metric is the arithmetic mean of MUC, B^3 and $CEAF_e$ metrics.

of 18.7 points, from 53.72 with automatic mentions to 72.42 when gold mentions are used.

We also had a look at the pronoun resolution performance alone, but only MUC scores on automatic mentions as the CoNLL scorer does not provide a break-down of scores per anaphor type, and there was a small gain in performance from the Soon baseline to the Basque model from $F1 = 27.4$ to $F1 = 33.0$, respectively. The gain is due mostly to higher precision, suggesting the additional features in the Basque model help discriminate better erroneously resolved pronouns in the baseline model, however, more work will need to be devoted to improving recall, which is particularly challenging in the case of Basque due to the lack of gender in the Basque pronoun system.

5.1 Error Analysis

In our error analysis we had a look at examples from our corpus covering the following four cases:

Case a. There were errors in the coreference resolution due to errors in the pre-processing which were propagated across the pipeline. Consider example 1, for instance:⁶

- (1) **Gold mentions:** [Del Bosque] prentsurrekoa eman zuen atzo. [Vicente Del Bosque], [Real Madrileko entrenatzailea], nahikoa kezkatu azaldu zen.
Automatic mentions: [Del Bosque] prentsurrekoa eman zuen atzo. [Vicente Del Bosque], [Real Madrileko entrenatzailea], nahikoa kezkatu azaldu zen.

Case b. Due to the challenges posed by the genderless pronoun system in Basque, there were pro-

⁶English translation: “[Del Bosque] gave a press conference yesterday. [Vicente Del Bosque], [Real Madrid coach], appeared quite concerned”.

nouns easy to resolve in relative terms which were missed or incorrectly resolved. Example 2 illustrates this:⁷

- (2) Lehendakari hautatu zutenetik, [Djukanovick] aldaketa handia eman dio [bere] ildo politikoari.

Case c. Here with example 3 we illustrate an instance of a challenging cases of pronouns which are currently beyond the scope of our approach:⁸

- (3) Gobernuaren bilera honen ondoren, oporretara joango da [Jospin], eta hauek baliatuko ditu, ziur aski, Chevenement kasuz gogoetak egiteko eta konponbide batekin [bere] jarduerari eusteko.

In this example it is more challenging to resolve correctly the pronoun [bere] “[his]” as [bere] can refer to Jospin or to Chevenement.

Case d. Finally, with example 4 we show an instance of a correctly resolved pronoun by our system:⁹

- (4) “[Guk] ez dugu inoiz penaltietan irabazi.” Luzapena golik gabe amaitzean, itzal beltz batek estali zuen Arena estadioa. Rijkaard-ek esana zuen arreta bereziz prestatu zituztela penaltiak, “[gure] istoria ez errepikatzeke”.

5.2 Discussion

Taking into consideration Basque most relevant grammatical characteristics, in some aspects it is more challenging to resolve coreferences in this language than in others.

Since Basque is an agglutinative language, a given lemma takes many different word forms, depending on the case (genitive, locative, etc.) or the number (singular, plural, indefinite) for nouns and adjectives. For example, the lemma *lehendakari* (“president”) forms the inflections *lehendakaria* (“the president”), *lehendakariak* (“the president”), *lehendakariari* (“to the president”), *lehendakariei* (“to the presidents”), *lehendakariaren* (“of the president”), etc. This means that looking only for the given exact word, is not enough for Basque

⁷English translation: “Since he was elected as president, [Djukanovic] has greatly changed [his] policy lines”.

⁸English translation: “After this government meeting, [Jospin] will go on holidays, and will surely use it to reflect on Chevenement case and to maintain [his] activity with a new solution”.

⁹English translation: “[We] have never won on penalties.” After the extension finished without goals, a large shadow turn off the stadium. Rijkaard said they prepared penalties with great attention, “so that [our] story would not occur again”.

to resolve coreference when string matching techniques are applied and as we observed in our experiments the use of lemmas is more effective in morphologically rich languages.

Besides the agglutination, there is no grammatical gender in the nominal system. Nouns and adjectives have no distinct endings depending on gender. In addition, there are no distinct forms for third person pronouns in Basque, and demonstratives are used as third person pronominals (Laka, 1996).

This makes it impossible to use gender as a feature in the resolution process which has been proven particularly useful in the resolution of pronouns, for example. Furthermore, the animacy feature cannot be used for pronoun resolution either. In this scenario, distance-based features, like Sentence Distance and Markable distance could be the most effective features for pronoun resolution. Nevertheless, research will have to be devoted to finding other useful features to make up for the lack of gender and animacy.

6 Conclusion

In this paper we presented our ongoing work on Coreference Resolution in Basque. We described the main resource we have been using which is the EPEC corpus annotated with coreferences and we explained how we have been adapting the coreference resolution toolkit, BART, to enable it to process Basque. We ran two levels of experiments one resolving coreferences using the gold mentions and one using automatically parsed mentions and we trained two different models for each, a baseline model based on (Soon et al., 2001) and a Basque model with extended feature set. We showed that the Basque model significantly outperforms the baseline. We also discussed key characteristics of the Basque language which make it particularly challenging for coreference.

Next we plan to investigate more in depth suitable features that can both make up for the lack of gender and animacy and be extracted reliably from unrestricted text. We also plan to run an extrinsic evaluation gauging the effect of coreference on a higher level task.

Acknowledgments

This work has been supported by Ander Soraluze's PhD grant from Euskara Errektoreordetza, the University of the Basque Country (UPV/EHU) and by the EXTRECM project, Spanish Government (TIN2013-46616-C2-1-R). The research leading to these results has received funding from the European Union - Seventh Framework Programme (FP7/2007-2013) under grant agreement 610916 SENSEI.

References

- Itziar Aduriz and Arantza Díaz de Ilarraza. 2003. Morphosyntactic Disambiguation and Shallow Parsing in Computational Processing of Basque. *Inquiries into the lexicon-syntax relations in Basque*, pages 1–21. University of the Basque Country.
- Itziar Aduriz, Maxux Aranzabe, Jose Mari Arriola, Maite Atutxa, Arantza Díaz de Ilarraza, Nerea Ezeiza, Koldo Gojenola, Maite Oronoz, Aitor Soroa, and Ruben Urizar. 2006. Methodology and Steps towards the Construction of EPEC, a Corpus of Written Basque Tagged at Morphological and Syntactic Levels for the Automatic Processing. pages 1–15. Rodopi. Book series: Language and Computers.
- Iñaki Alegria, Xabier Artola, Kepa Sarasola, and Miriam Urkia. 1996. Automatic Morphological Analysis of Basque. *Literary & Linguistic Computing*, 11(4):193–203.
- Iñaki Alegria, Maxux Aranzabe, Aitzol Ezeiza, Nerea Ezeiza, and Ruben Urizar. 2002. Robustness and Customisation in an Analyser/Lemmatiser for Basque. In *LREC-2002 Customizing knowledge in NLP applications workshop, pages 1-6, Las Palmas de Gran Canaria, 28th May 2002*.
- Iñaki Alegria, Nerea Ezeiza, Izaskun Fernandez, and Ruben Urizar. 2003. Named Entity Recognition and Classification for texts in Basque. In *II Jornadas de Tratamiento y Recuperación de Información*, (JOTRI 2003), pages 198–203, Madrid, Spain.
- Iñaki Alegria, Olatz Ansa, Xabier Artola, Nerea Ezeiza, Kepa Gojenola, and Ruben Urizar. 2004. Representation and Treatment of Multiword Expressions in Basque. In *ACL workshop on Multiword Expressions*, pages 48–55.
- Iñaki Alegria, Bertol Arrieta, Xavier Carreras, Arantza Díaz de Ilarraza, and Larraitz Uria. 2008. Chunk and Clause Identification for Basque by Filtering and Ranking with Perceptrons. *Procesamiento del Lenguaje Natural*, 41.
- Amit Bagga and Breck Baldwin. 1998. Algorithms for Scoring Coreference Chains. In *In The First International Conference on Language Resources and Evaluation Workshop on Linguistics Coreference*, pages 563–566.
- Samuel Broscheit, Simone Paolok Ponzetto, Yannick Versley, and Massimo Poesio. 2010. Extending BART to provide a coreference resolution system for German. In *Proceedings of the International Conference on Language Resources and Evaluation, LREC 2010*, Valletta, Malta.
- Klara Ceberio, Itziar Aduriz, Arantza Díaz de Ilarraza, and Ines Garcia-Azkoaga. 2016. Coreferential relations in Basque: the annotation process. Theoretical Developments in Hispanic Linguistics. The Ohio State University.
- Defense Advanced Research Projects Agency. 1995. *Proceedings of the Sixth Message Understanding Conference (MUC-6)*, San Francisco, CA. Morgan Kaufmann.
- Lynette Hirschman. 1998. MUC-7 coreference task definition, version 3.0. In N. Chinchor, editor, *Proceedings of the 7th Message Understanding Conference*. NIST.
- Mateusz Kopeć and Maciej Ogrodniczuk. 2012. Creating a Coreference Resolution System for Polish. In *Proceedings of the Eight International Conference on Language Resources and Evaluation (LREC'12)*, Istanbul, Turkey. European Language Resources Association (ELRA).
- Itziar Laka. 1996. A Brief Grammar of Euskara, the Basque Language. <http://www.ehu.es/grammar>. University of the Basque Country.
- Heeyoung Lee, Angel Chang, Yves Peirsman, Nathanael Chambers, Mihai Surdeanu, and Dan Jurafsky. 2013. Deterministic Coreference Resolution Based on Entity-centric, Precision-ranked Rules. *Comput. Linguist.*, 39(4):885–916, December.
- Xiaoqiang Luo. 2005. On Coreference Resolution Performance Metrics. In *Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing, HLT '05*, pages 25–32, Stroudsburg, PA, USA. Association for Computational Linguistics.
- Tom Morton. 2000. Coreference for NLP applications. In *Proceedings of the Annual Meeting of the Association for Computational Linguistics (ACL)*.
- Christoph Müller and Michael Strube. 2006. Multi-level Annotation of Linguistic Data with MMAX2. In Sabine Braun, Kurt Kohn, and Joybrato Mukherjee, editors, *Corpus Technology and Language Pedagogy: New Resources, New Tools, New Methods*, pages 197–214. Peter Lang, Frankfurt a.M., Germany.

- Massimo Poesio, Olga Uryupina, and Yannick Versley. 2010. Creating a coreference resolution system for Italian. In *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)*, Valletta, Malta, may. European Language Resources Association (ELRA).
- Massimo Poesio, Roland Stuckardt, and Yannick Versley, editors. 2016. *Anaphora Resolution: Algorithms, Resources and Applications*. Springer-Verlag.
- Sameer Pradhan, Xiaoqiang Luo, Marta Recasens, Eduard Hovy, Vincent Ng, and Michael Strube. 2014. Scoring Coreference Partitions of Predicted Mentions: A Reference Implementation. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 30–35. Association for Computational Linguistics.
- Marta Recasens and Eduard Hovy. 2011. BLANC: Implementing the Rand index for coreference evaluation. *Natural Language Engineering*, 17(4):485–510.
- Wee Meng Soon, Hwee Tou Ng, and Daniel Chung Yong Lim. 2001. A Machine Learning Approach to Coreference Resolution of Noun Phrases. *Computational Linguistics*, 27(4):521–544, December.
- Ander Soraluze, Olatz Arregi, Xabier Arregi, Klara Ceborio, and Arantza Díaz de Ilarraza. 2012. Mention Detection: First Steps in the Development of a Basque Coreference Resolution System. In *KONVENS 2012, The 11th Conference on Natural Language Processing*, Vienna, Austria.
- Ander Soraluze, Olatz Arregi, Xabier Arregi, and Arantza Díaz de Ilarraza. 2015. Coreference Resolution for Morphologically Rich Languages. Adaptation of the Stanford System to Basque. *Procesamiento del Lenguaje Natural*, 55:23–30.
- Josef Steinberger, Massimo Poesio, Mijail A. Kabadjov, and Karel Ježek. 2007. Two uses of anaphora resolution in summarization. *Information Processing and Management*. Special Issue on Text Summarisation.
- Olga Uryupina, Alessandro Moschitti, and Massimo Poesio. 2012. BART goes multilingual: the UniTN/Essex submission to the CoNLL-2012 shared task. In *Proceedings of the Joint Conference on EMNLP and CoNLL-Shared Task*, Jeju, Korea.
- Yannick Versley, Simone Paolo Ponzetto, Massimo Poesio, Vladimir Eidelman, Alan Jern, Jason Smith, Xiaofeng Yang, and Alessandro Moschitti. 2008. BART: a modular toolkit for coreference resolution. In *Proceedings of the 2008 Conference of the Association for Computational Linguistics*, pages 9–12.
- Marc Vilain, John Burger, John Aberdeen, Dennis Connolly, and Lynette Hirschman. 1995. A Model-theoretic Coreference Scoring Scheme. In *Proceedings of the 6th Conference on Message Understanding*, MUC6 '95, pages 45–52, Stroudsburg, PA, USA. Association for Computational Linguistics.