Dictionary and Monolingual Corpus-based Query Translation for Basque-English CLIR

Xabier Saralegi, Maddalen Lopez de Lacalle

R&D Elhuyar Foundation

Usurbil, Spain

x.saralegi@elhuyar.com, m.lopezdelacalle@elhuyar.com

Abstract

This paper deals with the main problems that arise in the query translation process in dictionary-based Cross-lingual Information Retrieval (CLIR): translation selection, presence of Out-Of-Vocabulary (OOV) terms and translation of Multi-Word Expressions (MWE). We analyse to what extent each problem affects the retrieval performance for the Basque-English pair of languages, and the improvement obtained when using parallel corpora free methods to address them. To tackle the translation selection problem we provide novel extensions of an already existing monolingual target co-occurrence-based method, the Out-Of-Vocabulary terms are dealt with by means of a cognate detection-based method and finally, for the Multi-Word Expression translation problem, a naïve matching technique is applied. The error analysis shows significant differences in the deterioration of the performance depending on the problem, in terms of Mean Average Precision (MAP), the translation selection problem being the cause of most of the errors. Otherwise, the proposed combined strategy shows a good performance to tackle the three above-mentioned main problems.

1. Introduction

CLIR is becoming an increasingly relevant topic due to the growth in multilingual information and the fact that most inhabitants are polyglots. A typical CLIR system offers the user searching topics in his or her mother tongue and retrieves documents in other languages. Different strategies exist to tackle the crosslinguality depending on what information is translated: topics, documents or both. The best results are obtained by translating the collections into the language of the queries. However, this approach is computationally expensive and most of the works have focused on query translation methods. These methods can be based on MT systems, parallel corpora or dictionaries. MT systems and parallel corpora are scarce for the majority of language pairs. That is why we think that the dictionary-based query translation approach must be explored, since bilingual dictionaries are more abundant and easier to obtain. That is the circumstance of the Basque-English language pair. In the dictionary-based query translation task well-known problems arise that need to be solved, some of the most relevant being translation selection, presence of OOV terms and MWE translations. We propose methods based on target co-occurrences to deal with translation selection, cognate detection to deal with OOV terms, and a naïve matching process to detect MWEs. It is important to notice that all the methods presented in this paper are parallel corpora free. In addition to addressing these problems, we are also interested in measuring exactly how each problem affects retrieval performance in dictionary-based query translation and how good the proposed methods deal with them. We need a gold standard to do that evaluation. So we detect and fix the aforementioned three problems manually, and we consider this to be the reference theoretical optimum or topline performance of the system.

The paper is organized as follows: first, we review some related works in which different methods to treat inherent problems in CLIR are presented. Next, the strategy we are proposing for translating the query is introduced, along with the methods it involves. That is followed by an appraisal of how each problem affects retrieval performance and how well the proposed methods tackle it. Finally, evaluation results and conclusions are presented.

2. Related Work

CLIR can be seen as IR with a language barrier placed between the query and the collection. Even though most authors choose to translate the queries into the language of the target collection, mainly due to the lower requirements of memory and processing resources (Hull and Grefenstette, 1998), documents have richer context information than queries, are useful in the translation selection process, and have more examples to reduce error rate of translations. (Oard, 1998) proved that under certain conditions the quality of the translation and retrieval performance improve when the collection is translated. Furthermore, translating both queries and documents and merging the obtained ranks provides even better results (McCarley, 1999; Chen and Gey, 2003). The different techniques to carry out the translation can be grouped as follows, depending on the translation-knowledge source: MT-based, parallel corpus-based, and bilingual dictionary-based. For the last two groups different statistical frameworks are proposed; cross-lingual probabilistic relevance models and cross-lingual language models. The first one offers useful operators to treat the ambiguous translations and is usually used along with dictionaries. The second one incorporates translation probabilities on a more formal and unified framework which are obtained from parallel

corpora (Hiemstra, 2000)). The results depend on the quality of the resources but usually better results are achieved with cross-lingual language models (Xu et al., 2001).

However, parallel corpora are a scarce resource. Dictionaries are more accessible but the ambiguous translations must be dealt with. For the translation selection (A. Pirkola, 1998) proposed to use structured queries along with probabilistic relevance models. In this approach all translations of a source word are treated as the same token when TF and DF statistics are calculated for the translations of that source word. (Darwish and Oard, 2003) introduce a probabilistic structured query where weights are applied to translation candidates when TF and DF values are calculated. It offers improvement over non-probabilistic structured queries but only when parallel corpora are used to estimate the weights. As an alternative, (Saralegi and Lopez de Lacalle, 2010) proposed that these weights be estimated by calculating the cross lingual distributional similarity between contexts of the translation candidates obtained from the web, using the web as a comparable corpus.

Other authors propose using the target collection as a language model to solve the translation selection problem (Monz and Dorr, 2005; Ballesteros and Croft, 1998; Gao et al., 2001). The proposed algorithms try to select the translation candidates which show the highest association degree in the target collection. The algorithms differ in the way the global association is calculated and in the translation unit used (i.e., word, noun phrases...) (Monz and Dorr, 2005; Gao et al., 2001;Gao et al., 2002; Liu et al., 2005).

Structured queries and co-occurrences-based methods were compared in (Saralegi and Lopez de Lacalle, 2009). There was no significant difference in results when dealing with short queries. But when dealing with long queries, structured queries offer a significantly better MAP than the co-occurrences-based method. This is probably due to the synonym expansion effect produced and the implicit retrieval time selection, which is better when a long context is provided.

The other main problems which affect the translation process are the presence of OOV terms and the translations of MWEs. Cognate detection is the main strategy used for OOV terms treatment (Knight and Graehl, 1997).The translation of the MWE is also explored in some papers (Ballesteros and Croft, 1997).

3. Proposed Query Translation Method

In this work we have designed a global method that combines state-of-the-art and novel techniques to tackle the aforementioned problems in query translation. We propose a cognate detection-based method to find the translations of the OOV words in the target collection. To address the translation selection problem we propose a target co-occurrences-based method, based on the one proposed by (Monz and Dorr, 2005). Although this method did not obtain better results compared with the ones obtained with structured queries in previous works (Saralegi and Lopez de Lacalle, 2009), the truth is that the syn operator of the structured queries is not provided by all retrieval models. Hence, we carried out our experiments with the co-occurrence-based approach. For the MWE treatment we used a simple matching and

translation technique based on a bilingual MWE list to detect and translate them.

3.1 Experimental Setup

We prepared two sets of topics: a set of topics belonging to the CLEF 2001 edition (41-90) that was used as the development set, and another set of topics for test purposes (250-350). All topics were translated by hand from English to Basque. These topics were lemmatized in both languages. We also used the corresponding collections and human relevance judgements. It must be noted that only the LA Times 94 collection is related to the queries of the development set whereas both LA Times 94 and Glasgow Herald collections are linked to the test queries. We adopted a dictionary-based method to carry out the translation process. We used the *Morris* Basque/English dictionary including 77,864 entries and 28,874 unique Basque terms, and the *Euskalterm* terminology bank including 72,184 entries and 56,745 unique Basque terms. According to (Demner-Fushman and Oard, 2003) the growth in mean average precision is evident between about 3,000 and 20,000 unique terms. They conclude that beyond that range, little further improvement is observed. Hence, we can assume that the coverage of our dictionary is sufficient for the query translation task. We used the *Indri* retrieval algorithm for all the runs.

3.2 Treating Out-Of-Vocabulary words

The proposed cognate detection approach consists of applying some transliteration rules to the OOV word and then looking for its cognates in the target collection, by computing the Longest Common Subsequence Ratio (LCSR) measure between the transliterated OOV word and words in the target collection..

In order to measure the damage caused by OOV words in the translation and retrieval processes, we first quantified out these kinds of words in the development set of topics. A total of 64 OOV terms were quantified out and they account for 15.46% of all query terms. This is a normal number taking into account the size of our dictionary. Afterwards, we determined the number of OOV words translated correctly by applying cognate detection. There were 89% in all, and almost all of them were named entities like in the study carried out by (Demner-Fushman and Oard. 2003). Despite the fact that this was a good result, we realized that only a total of 7 (10.94%) OOV words needed transliteration and LCSR to detect their translation (Examples in Table 1). The rest of the resolved OOV words were named entities and words that are written equally in both languages. We classified the OOV words depending on their POS. (See Table 2).

OOV word	Tran rule	Trans-literation	Max. LCSR
txetxenia	tx/ch	chechenia	(chechenia,chechenya) =0.89
korrupzio	-zio/ -tion k/c	corruption	(corruption,corruption) =1

Table 1. Example of an OOV word resolved using cognate detection

Named Entities	Nouns	Adj.	Numbers
82.81%	12.5%	3.13%	1.56%

Table 2 Distribution of OOV words depending on their POS

We can see that even if the number of OOV words resolved with the cognate detection-based method are only a few, with respect to the MAP value, using the cognate detection-based method was effective (See Table 3). So, it seems that OOV words tend to be relevant terms in the query, named entities in their majority. We translated OOV words by hand and calculated the MAP value to estimate the topline. The fall produced when OOV words are not treated is 4-12% (First translation of the dictionary or the OOV word itself). But after the proposed method is applied, the fall is reduced to 0.58-3.4%.

Translation method	MAP			
	Title		Title+Descrip.	
		Impr. Over First. %		Impr. Over First. %
First translation	0.2703		0.3835	
First translation + OOV(by hand)	0.3085	12.38	0.3999	4.101
First translation+cognates	0.2969	8.96	0.3975	3.52

Table 3. Retrieval performance for OOV words for 41-90 topics

3.3 Translating Multi-Word Expressions

We identified the MWEs in the development set of topics by hand and analyzed whether they were compositional, in other words, whether they could be translated word by word or not. A total of 60 MWEs were quantified out and exactly 52 (%86.67) of them could be translated word by word (Example on Table 4).

Basque MWT	Words	Translations from Dictionary	Correct Candidate
bigarren mundu gerra	bigarren	second, secondary	second
	mundu	people, world	world
	gerra	war	war

Table 4. Example of word-by-word MWT translation

We compared retrieval performance by taking the first translation of each word in the MWE and taking the translation of the complete MWE from the dictionary when available. In addition, we translated all the MWEs by hand and calculated the MAP in order to estimate the topline resulting from the treatment of all of the MWEs. A total of 11 MWEs were directly translated from the dictionary. However, only the translations for the Basque

MWE *"esku hartze"* and *"eguzki energia"* Basque MWE translations differ from the ones obtained with the word by word translation. Although they are very few, it seems they tend to be relevant, as a significant improvement is achieved in terms of MAP (See Table 5). The proposed terminology list-based matching method does not offer a good result, maybe due to its dependence on the recall of the terminology bank. However, as the majority of MWEs are compositional, the co-occurrence-based translation selection method solves most of them.

Translation method	MAP			
	Title		Title+Descrip	
		Impr. Over First. %		Impr. Over First. %
First translation	0.2703		0.3835	
First translation+MWE (by hand)	0.3371	19.81	0.4222	9.17
First translation+MWE	0.2860	5.49	0.3944	2.76

Table 5. Retrieval performance for MWEs for 41-90 topics

3.4 Translation selection based on target co-occurrences

Finally, we proposed an algorithm based on target collection co-occurrences to deal with the translation selection problem. We adopted the implementation proposed by (Monz and Dorr, 2005):
Initially, all the translation candidates are equally likely. Assuming that t is a translation candidate of the set of all candidates $tr(s_i)$ for a query word s_i given by the dictionary, then:
Initialization step:

$$w_T^0(t/s_i) = \frac{1}{|tr(s_i)|}$$

In the iteration step, each translation candidate is iteratively updated using the weights of the rest of the candidates and the weight of the links connecting them.
Iteration step:

$$w_T^n(t/s_i) = w_T^{n-1}(t/s_i) + \sum_{t' \in inlink(t)} w_L(t,t') \cdot w_T^{n-1}(t'/s_i)$$

where $inlink(t)$ is the set of translation candidates that are linked to t, and $w_L(t,t')$ is the association degree between t and t' in the target collection, measured by the log-likelihood ratio.
After re-computing each translation candidate weight, they are normalized.

Normalization step:

$$w_T^n(t/s_i) = \frac{w_T^n(t/s_i)}{\displaystyle\sum_{m=1}^{|tr(s_i)|} w_T^n(t_{i,m}/s_i)}$$

The iteration stops when the variations of the term weights become smaller than a predefined threshold.

In order to measure how the translation selection problem affects retrieval performance we set up two toplines. One involved selecting the correct translation from among those candidates given by the dictionary by hand; in the other, a new translation was also provided if it was not in the dictionary. This new translation was taken from the corresponding source English query. We saw that the MAP results obtained in both experiments were notably better than those obtained with the baseline (First translation) (See Table 5). However, it is noteworthy that introducing new translations outperforms the method including the hand selected translations only. Hence, rather than a selection problem, it would depend on the translation recall of the dictionary used. So for this system the topline will be determined by "Translation selection by hand" results. The Monz and Dorr selection algorithm (Target co-occurrence-based) achieves very similar results.

3.4.1 Adding a nearness factor to the degree of association

We introduced a variant into the Monz and Dorr algorithm. We modified the iteration step by adding a factor $w_F(t,t')$ to increase the association degree $w_L(t,t')$ between translation candidates t and t' whose corresponding source words $so(t), so(t')$ are near each other in the source query Q, and belong to the same MWE.

$$w'_L(t,t') = w_L(t,t') \cdot w_F(t,t')$$

$$w_F(t,t') = \frac{\max\limits_{s_i,s_j \in Q} dis(s_i,s_j)}{dis(so(t),so(t'))} \cdot 2^{smw(so(t),so(t'))}$$

$$smw(s,s') = \begin{cases} 1 & \{s,s'\} \subseteq Z \ where\, Z \in MWU \\ 0 \end{cases}$$

According to the MAP scores, the proposed variant (Target co-occurrences-based + nearness) does not achieve any improvement (See Table 6).

Translation method	MAP			
	Title		**Title+Description**	
		Impr. Over First. %		**Impr. Over First. %**
First translation	0.2703		0.3835	

Translation selection by hand	0.3430	21.19	0.4266	10.10
Target co-occurrence based	0.3405	20.62	0.4123	6.99
Translation selection by hand+new translations	0.4004	32.49	0.4593	16.50
Target co-occurrence based+nearness	0.3399	20.48	0.4117	6.85

Table 6. Retrieval performance for translation selection for 41-90 topics

3.4.2 Calculating co-occurrences of senses instead of tokens

We also implemented another variant of the target collection co-occurrence-based algorithm, which instead of measuring the degree of association between the customary translation candidate words, it, measures the degree of association between the senses of the translations.

For example, for the source query word s_1 (e.g., metro) the senses of translations in the dictionary are C_1 and C_2, whose translation candidates are t_1 and t_2 (e.g., underground and subway) for the sense C_1 and t_3 and t_4 (e.g., metre and meter) for the sense C_2, $tr(s_1)=\{\{t_1,t_2\}\{t_3,t_4\}\}=\{C_1,C_2\}$. In the same way, the translation candidate for the source query word s_2 (e.g., geltoki) is t_5 (e.g., station) which belongs to the same and unique sense c_3, $tr(S_2)=\{\{t_5\}\}=\{C_3\}$. Thus, the frequency for a sense will be calculated as the sum of the frequencies of all the translation candidate words that belong to that sense. Continuing with the example, the frequency of the sense C_1 will be calculated as the sum of the frequencies of the words t_1 and t_2, $f(C_1)=\sum_{t\in C_1} f(t)$ and the frequency of the sense C_2 as the frequency of the words t_3 and t_4, $f(C_2)=\sum_{t\in C_2} f(t)$ and lastly, the frequency of the sense C_3 as the sum of the frequency of the word t_5, $f(C_3)=\sum_{t\in C_3} f(t)$. Thus, the frequency with which the senses c_1 and c_3 appear together in the same document will be calculated as the intersection of the union of the translation candidates belonging to each sense $f(C_1 \cap C_3)=f((\bigcup_{t\in C_1} t) \bigcap (\bigcup_{t\in C_3} t))$.

In order to compute $f(C_1 \cap C_3)$ faster, we built a new target collection which contained the senses of the words. The tokens of the collection will be formed by joining the corresponding source word and the sense taken from the dictionary (*source_word_id + sense_id*). So if a translation appears in more than one dictionary entry, all the senses will be taken for the new collection by introducing as many new tokens as senses where it appears.

The results show that in the case of long queries the new method offers a significant MAP improvement over the Monz and Dorr algorithm (target co-occurrence-based). (See Table 7).

Translation method	MAP			
	Title		Title+Description	
		Impr. over First %		Impr. over First %
First translation	0.2703		0.3835	
Target token co-occurrence based	0.3405	23.29	0.4059	5.52
Target sense co-occurrence based	0.3323	18.05	0.4163	7.88

Table 7. Retrieval performance for sense-based translation selection for 41-90 topics

4. Evaluation

In order to carry out the evaluation we used a new set of topics belonging to the CLEF 2001 edition (250-350) and then used the corresponding collection (LA Times 94 and Glasgow Herald 95) and human relevance judgements.

First, we evaluated each of the proposed methods to deal with the problems in the Basque to English query translation task. Then, we evaluated different combinations of all methods:

- English monolingual or topline
- Baseline: Taking the first translation from the dictionary.
- OOV: First sense from the dictionary and cognate detection-based method to deal with OOV.
- MWE: MWE matching and first sense from the dictionary.
- Monz: Co-occurrence-based selection.
- Monz+Nearness: Co-occurrence-based selection including the nearness factor.
- Monz (senses): Sense co-occurrence-based selection.
- Monz (senses)+OOV: Sense co-occurrence-based selection and cognate detection-based method to deal with the OOV problem.

Translation method	MAP					
	Title			Title+Description		
		% over Mon.	Impr over First %		% over Mon	Impr over First %
English monolin.	0.3176			0.3773		
Baseline	0.2195	67		0.2599	69	
OOV	0.2279	72	7.24	0.2670	71	2.66
MWE	0.2237	70	5.5	0.2601	69	0.08
Monz	0.2315	73	8.68	0.2642	70	1.63
Monz-Nearness	0.2318	73	8.8	0.2627	70	1.07
Monz (senses)	0.2362*	74	10.5	0.2747	73	5.39
Monz (senses)+OOV	0.2424*	76	12.79	0.2805	74	7.34

Table 8. MAP values for 250-350 topics

In the results obtained (See Table 8), we can see that the translation selection problem is the one which is better dealt with. The co-occurrence-based translation selection significantly outperforms the first translation approach when dealing with short queries. The improvement offered by the co-occurrence-based method for long queries is lower, probably because the first translation method achieves better results when queries provide many terms. In addition, this lower improvement may be caused by the greedy nature of the translation selection algorithm. Since it has to deal with more translation candidates, it is more likely to reach a maximum. On the other hand, the new proposed sense co-occurrence -based extension exceeds the MAP value obtained with Monz and Dorr algorithm. Otherwise, as we have seen in the development experiments, the matching method proposed to deal with MWE translations offers a very poor performance. On the contrary, a cognate-based method for treating OOV words seems to be adequate. The best results are achieved by combining the sense co-occurrence-based translation selection method and the cognate-based OOV term translation method.

The improvements that are statistically significant according to the Paired Randomization Test with α=0.05 are marked with an asterisk in table 7.

5. Conclusions

We have developed a query translation method which tackles three main problems in dictionary-based CLIR: presence of OOV words, translation of MWEs, and treatment of ambiguous translations. We have analyzed how each problem affects the retrieval performance in terms of MAP. Although results change depending on the length of the queries, the decrease produced by the translation selection (10-21% drop) and the one produced by MWEs (9-20% drop) seem to be the more determining ones. In the case of translation selection, we can distinguish two cases: wrong selection from the dictionary (10-21% drop), and incorrect translations in the dictionary (17-32% drop). OOV treatment (4-12% drop) seems to be the least influential factor, probably due to the similar orthography of both languages. Other pieces dealing with evaluation issues of errors, derived from the MT-based translation process have been carried out (Zhu and Wang, 2006; Qu et. al, 2000). (Qu et al., 2000) point out that the wrong translation selection is the most frequent error in an MT-Based translation process. The same conclusion is obtained from our tests. In the development experiments, we have seen that the proposed methods for treating OOV words and ambiguous translations offer a good performance. The matching method proposed to treat MWEs offers a poor performance, but taking into account that almost all of

the MWEs are compositional, it is to be expected that they will be properly addressed by the co-occurrence-based translation selection method.

Otherwise, the improvements developed over the co-occurrence-based translation selection show a different performance behavior. Including the nearness factor provides a few better translations but this leads to no improvement in the overall retrieval performance. For example, the Basque query *"Antarktika balea ehiza debekatu"* is translated as *"Antarctic whale **hunting** forbidden"* adding the nearness factor while Monz and Dorr algorithm provides a slightly worse translation: *"Antarctic whale **game** forbidden"* Calculating co-occurrences between senses by means of the proposed method instead of between tokens provides better translation quality as well as better retrieval performance.

1. References

Ballesteros, L. and Croft, W.B. (1997). Phrasal translation and query expansion techniques for cross-language information retrieval. In *Proceedings of the 20th annual international ACM SIGIR conference on Research and development in information retrieval*.

Chen, A. and Gey, F.C. (2003). Combining Query Translation and Document Translation in Cross-Language Retrieval. In *Proceedings of the 4th Workshop of the Cross-Language Evaluation Forum*, pp. 108--121.

Darwish, K. and Oard, D.W. (2003). Probabilistic structured query methods. *In Proceedings of the 26th annual international ACM SIGIR conference on Research and development in information retrieval*, pp. 338–344.

Demner-Fushman, D. and Oard, W. D. (2003). The effect of bilingual term list size on dictionary-based cross-language. In *Proceedings of the 36th Annual Hawaii International Conference on System Sciences (HICSS'03).*

Gao, J., Nie, J.Y., Xun, E., Zhang, J., Zhou, M. and Huang, C. (2001). Improving Query Translation for Cross-language Information Retrieval using Statistical Models. In *Proceedings of the 24th annual international ACM SIGIR conference on Research an development in information retrieval*, pp. 96-104.

Gao, J., Nie, J.Y., He, H., Chen, W. and Zhou, M. (2002). Resolving query ambiguity using a decaying co-occurrence model and syntactic dependence relations. In *Proceedings of the 25th annual international ACM SIGIR conference on Research and development in information retrieval*, pp.183–190.

Hiemstra, D. (2000). Using language models for information retrieval. Doctoral Thesis. University of Twente.

Hull, D.A. and Grefenstette, G. (1998). Querying Across Languages: A Dictionary-Based Approach to Multilingual Information Retrieval. In *Proceedings of the 19th annual international ACM SIGIR conference on Research and development in information retrieval*, pp. 49--57.

Knight, K. and Graehl, J. (1997). Machine transliteration. In *Proceedings of the 35th Annual Meeting of the Association for Computational Linguistics*, pp. 128--135.

Liu, Y., Jin, R. and Chai, J.Y. (2005). A maximum coherence model for dictionary-based cross-language information retrieval. In *Proceedings of the 28th annual international ACM SIGIR conference on Research and development in information retrieval*, pp. 536-543.

McCarley, J. S. (1999). Should we translate the documents or the queries in cross-language information retrieval?. In *Proceedings of the 37th annual meeting of the Association for Computational Linguistics on Computational Linguistics*, pp. 208--214.

Monz, C. and Dorr, B.J. (2005). Iterative translation disambiguation for cross-language Information Retrieval. In *Proceedings of the 28th annual international ACM SIGIR conference on Research and development in information retrieval*, pp. 520--52.

Oard, D. W. (1998). A Comparative Study of Query and Document Translation for Cross-Language Information Retrieval. In *Proceedings of the Third Conference of the Association for Machine translation in the Americas (AMTA) Philadelphia*, pp.208--214.

Pirkola, A. (1998). The effects of query structure and dictionary setups in dictionary-based cross-language information retrieval. In *Proceedings of the 21st annual international ACM SIGIR conference on Research and development in information retrieval*, pp. 55--63.

Xu, J., Weischedel, R. and Chanh, N (2001). Evaluating a probabilistic model for cross-lingual information retrieval. *Proceedings of the 24th annual international ACM SIGIR conference on Research and development in information retrieval*, pp. 105-11.

Saralegi, X. and Lopez de Lacalle, M. (2009). Comparing different approaches to treat Translation Ambiguity in CLIR: Structured Queries vs. Target Co-occurrence Based Selection. In *Proceedings of the* 6th International Workshop on Text-based Information Retrieval.

Saralegi, X. and Lopez de Lacalle, M. (2010). Estimating Translation Probabilities from the Web for Structured Queries. In *Proceedings of the 32nd European Conference on Information Retrieval.*

Zhu, J. and Wang, H. (2006). The Effect of Translation Quality in MT-Based Cross-Language Information Retrieval. In *Proceedings of the 21st International Conference on Computational Linguistics and 44th annual Meeting of the ACL.*