Sofia N. Galicia-Haro¹ and Alexander F. Gelbukh²

¹ Facultad de Ciencias Universidad Nacional Autónoma de México, México, D. F. sngh@fciencias.unam.mx

> ² Centro de Investigación en Computación Instituto Politécnico Nacional, México, D. F. gelbukh@gelbukh.com

Abstract. This paper reports research on Spanish temporal expressions. The analyzed phrases include a common temporal expression for a period of years reinforced by an adverb of time. We found that some of those phrases are age-related expressions. We analyzed a sample obtained from the Internet to determine the local context where the age-related meaning is unmistakable. We present the results for 21 selected classes.

Keywords: temporal expressions, age of persons, local context

1 Introduction

Some words or whole sequences of words in a text are temporal expressions: for example, *yesterday*, *Monday 12th*, *two months*, *about a year and a half*; each refers to a certain period of time. Such words or sequences of words mainly share a noun or an adverb of time: *yesterday*, *month*, *year*. This leads to the problem of automatically deciding whether a word or a sequence is a temporal expression. It is an important part of many natural language processing applications, such as question answering, machine translation, information retrieval, information extraction, text mining, etc., where robust handling of temporal expressions is necessary.

Automatic recognition of expressions of time was introduced in the Named Entity Recognition task of the Message Understanding Conferences¹, where temporal entities were tagged as "TIMEX". Since then, researchers have been developing temporal annotation schemes, for example [2] for English, [12] for the much more constrained domain of meeting scheduling.

The authors in [2] produced a guideline intended to support a variety of applications in the performance of some useful tasks. As the authors pointed out, the guideline was not intended to represent all the varieties of temporal information conveyed in natural language communication. They were interested in temporal expressions that

^{*} Work partially supported by Mexican Government (CONACyT, SNI, CGPI-IPN, PIFI-IPN)

¹ http://timexportal.wikidot.com/timexmuc6

reference calendar dates, times of day, or durations. They considered lexical triggers to identify the temporal expressions. A lexical trigger is a word or numeric expression whose meaning conveys a temporal unit or concept. To be a trigger, the referent must be able to be oriented on a timeline, or at least oriented with relation to a time (past, present, future).

In this work, we analyzed other different temporal expressions that do not fulfill the previous trigger characteristics. These phrases are recognized by an initial adverb: for example, *around*, *still*; and they end with a noun of time such as *year*, *month*. For example²: *aún en los últimos años* "still in the last years", *aún en los setentas* "still in the seventies", *alrededor de año y medio* "about a year and a half", *alrededor de hace 20 años* "about 20 years ago".

We found that this type of phrase presents interesting cases. There is for example a group that describes a person's age. Automatic recognition of a person's age should be useful in question answering and machine translation tasks, among others. For example, considering the following sentence:

El quíntuple retador mundialista Lazarte no claudica en su intención de coronarse campeón del mundo, aún a sus 38 años 'The five-time World Cup challenger Lazarte did no falter in his intention to be crowned world champion, although he is 38 years old'.

Spanish native speakers would understand that the phrase *aún a sus 38 años* denotes Lazarte's age. However, general machine translators give the wrong phrases 'yet his 38 years' or 'although his 38 years'. Also, this sentence would be retrieved by a question answering system to the specific question ¿How old is Lazarte?

In this article, we present a corpus-based analysis carried out to determine the context of such temporal expressions and their automatic determination. In section 2, we present the characteristics of the phrases we are interested in. In section 3 we describe the method we applied to obtain the materials for the analysis. We present the analysis of their local context in section 4. Finally, possible future work and conclusions are presented in sections 5 and 6.

2 Adverbs and Temporal Expressions

Adverbs of time (for example: *before, today, after, tomorrow*) create cohesion and coherence by forming time continuity in the events expressed in texts [8]. They are closely associated with narrative texts but they also appear in newspaper texts, for example: *Senators will approve today in the morning a schedule for the debate, after the plenary session the coordinators will meet to establish the regulation and tomorrow the tribune will be liberated.*

Researchers have been interested in adverbs of time included in temporal expressions, for example [2, 10, 11]. We chose to analyze the Spanish temporal phrases that begin with an adverb of time (AdvT) and end with a noun of time (TimeN), for example:

aún a principios de los años 90 "still at the beginning of the 1990s" aún en nuestros días "still in our days"

² The translations of the Spanish examples are literal.

aún hoy día, aún hoy en día "still nowadays"

alrededor de los 20 años "about 20 years"

These phrases present interesting issues. We can observe the relation between the groups of words in the following examples:

1. A sus 30 años Juan se comporta como niño

2. Aún a sus 30 años Juan se comporta como niño

3. Hoy a sus 30 años Juan se comporta como niño

The sentences describe the same main fact: *John, who is 30 years old, behaves like a child*, but they tell us something else when we introduce a modifier [1] (*aún* "still", *hoy* "today") in each one: they argue for different conclusions.

(1) Still, at 30 years old, John behaves like a child \Rightarrow in spite of his age he behaves as if he were a child

(2) Today, at 30 years old, John behaves like a child \Rightarrow today he behaves like a child

The adverbs "still" and "today" make such conclusions obligatory and reinforce the meaning of time in different forms. Both adverbs are related to time duration, one strict reading refers to 24 hours and the other to a longer period of time, but they also imply a direct judgment on the perception of the speaker, on the behavior of the subject or both. Finally, these conclusions are the main scientific objective of this work but here we began by trying to solve the initial step of automatic determination of such phrases.

For the analysis of temporal expressions reinforced by an adverb of time, we use one text collection compiled from a Mexican newspaper that is published daily on the Web in almost its entirety. The texts correspond to diverse sections, economy, politics, culture, sport, etc., from 1998 to 2002. The text collection has approximately 60 million words [3].

We wrote a program to extract the sentences matching the following pattern:

AdvT-somet	hing_TimeN
Auvi sonici	ming runciv

Where:

something –	corresponds to a sequence of up to six words' without punctuation
	marks, verbs or conjunctions.
TimeN –	corresponds to the following nouns of time: <i>año</i> "year", <i>mes</i>

- corresponds to the following nouns of time: *año* "year", *mes* "month", *día* "day", *hora* "hour", *minuto* "minute", *segundo* "second"

AdvT – adverbs of time, a collection of 51 elements from a dictionary⁴.

The extracted sentences were analyzed in [3]. From a subset of these sentences where the adverb corresponds to *actualmente* "at present", *ahora* "now", *alrededor* "around', *aún* "still', and the noun of time corresponds to *año* "year', we found some phrases expressing age of persons.

Usually the age of persons is described by Spanish temporal expressions including the time nouns *años* "years" and *meses* "months" (for babies). They can be recognized in the following ways:

³ A larger quantity of words does not guarantee any relation between the AdvT and the TimeN

⁴ DRAE, Real Academia Española. (1995): *Diccionario de la Real Academia Española*, 21 edición (CD-ROM), Espasa, Calpe

SEARCH(C)

For each phrase of type ADV-*-NounT or string-*-NounT in C

(1) Obtain 100 examples from the Internet

(1.1) $D = \{\text{examples excepting such instances where * includes verbs or punctuation}\}$

(1.2) Print D

(2) Classify them according to such words retrieved by *

(3) For each group of phrases sharing words retrieved by *, assign a class D_i

(3.1) F = class D_i

(3.2) SEARCH(F)

UNTIL no new elements are obtained

Fig. 1. Algorithm to obtain variants of temporal expressions

(1) with a string "de edad" ("old" in English) after the word años,

- (2) a number and years after the person's name, delimited by commas [9],
- (3) with the strings: *la edad de, de edad de* (lit. "the age of", "of the age of") before the number of years.

There are, however, other temporal expressions that describe the age of persons: for example, *aún a sus 65 años*, "still at his 65 years", *de alrededor de 20 años*, lit. "of about 20 years". These temporal phrases denote a point in the timeline of a person; it could be a point in the timeline of the events related in the sentence or a point in a tangential timeline.

We manually select one arbitrary example representing what we consider a class: a different combination of an adverb and a preposition before the number of years, the five resulting classes correspond to *aún a, aún con, actualmente de, alrededor de, ahora de.*

3 Searching for More Examples

Since our newspaper text collection contains a subset of all possible temporal phrases expressing the ages of persons, we analyzed a method to obtain a more representative group of phrases. Different possibilities exist for obtaining such variants, and we chose to look for examples on the Internet. This option allowed us to find phrases generated by native speakers more quickly, including the commoner collocations. Nevertheless, we know that searching the Internet has drawbacks, as [7] has already suggested. For example, the searching machines deliver different accounts for repetitions of the same question although this problem does not turn out to be so important in this work. One significant drawback for our work is that the results are classified in accordance with complex and unknown algorithms, so we cannot know what predispositions have been introduced for the order in which results are presented. Despite these considerations, we decided to search the Internet on the basis that we do not know how the results are classified.

The main idea of obtaining more examples from the Internet is based on getting a few examples from the newspaper texts (corresponding to the five classes above mentioned), simplifying them (eliminating determinants, adjectives, etc.) and searching for variants by including Google's asterisk facility [5]. The whole procedure is shown in Figure 1. For example: for the phrase *aún con sus jóvenes 48 años* the string when simplified becomes "*aún con año*" and the search is "*aún con * años*" using the Google search engine tool limited to the Spanish language where the asterisk substitutes for the eliminated words. Google returns hits where there is a string of words initiated by "*aún con*" then a sequence of words, ending with "*años*". The example for the whole procedure in Figure 1 is presented as follows:

SEARCH("aún con * años")

Step (1) 100 examples

... y el bachillerato en Lleida, *aún con dieciséis años* entró a trabajar de chico ... *aun con tantos años* sigo siendo el mismosde siempre...

... porque se dio cuenta que aún con tantos años encima son capaces de ...

Un partido, aún con pocos años de actuación, inspirado en la Gran ...

... y menos aún con Nuestros años felices (1996) o Tarde (1998)

Step (2) D= { aún con dieciséis años, aún con pocos años, aun con tantos años, aún con Nuestros años, ... }

Step (3) For each one of the classes a new process is initiated

SEARCH("aún con tantos * años") SEARCH("aún con pocos * años")

The process is repeated several times until no new repeated phrases are obtained, determining the sequences of words that appear with higher frequency.

We note that in addition some phrases not corresponding to the temporal phrases we are interested in are picked up: for example, *aún con tecnología 40 años más moderna*... "still with technology 40 years more modern", this type of phrase is eliminated in the manual identification at the end of the whole process.

After this compilation of examples, we manually select 21 classes that appear in the first column of Table 1, where NUM considers numbers represented by digits or letters.

4 Context for Phrases Denoting Age of Persons

We found that some of the 21 classes obtained from the Internet seem to preserve their meaning independently of the context and others require some form of words in context to denote the age of a person. To automatically determine the meaning of these temporal phrases assessing the context is the main issue.

To analyze the context required to preserve the person's age meaning we obtain examples by searching again on the Internet. The quantity of pages automatically obtained was limited to 50, i.e. to obtain 500 snippets.

Context information is considered in this work as two words in a window surrounding the target phrase without consideration of grammatical relations of the

whole sentence in an automatic form by means of software. We applied the system for automatic morphological analysis of Spanish [6] to assign parts of speech to the words of the context.

We wrote a program to classify the examples according to the context, to obtain nominal phrases, and to eliminate temporal phrases not matching the pattern AdvT-something-TimeN. We manually performed the general syntactic and semantic analysis of the context in the sentence.

Type of phrase	# examples	% age- related	# short snippet	# not age
aún a sus NUM años	293	96	7	5
aún hoy a sus NUM años	38	92	2	1
aún a tus NUM años	7	100	0	0
ahora a mis NUM años	182	99	1	1
aún con mis pocos NUM años	1	100	0	0
aún con mis cortos NUM años	2	100	0	0
aún con sus escasos NUM años	4	100	0	0
aún con tus casi NUM años	1	100	0	0
aún con sus casi NUM años	7	57	0	3
aún con sus NUM años	109	86	0	15
ahora de NUM años	352	86	6	45
de alrededor de NUM años	353	44	4	194
actualmente con NUM años	270	80	3	50
actualmente de NUM años	28	36	7	11
actualmente de unos NUM años	16	19	1	12
ahora con casi NUM años	118	67	0	39
ahora con más de NUM años	90	46	7	45
ahora a NUM años	112	0.9	8	103
ahora a los NUM años	132	36	2	82
alrededor de NUM años	242	16	9	194
alrededor de los NUM años	355	84	4	54

Table 1. Overall results for the examples obtained from the Internet

In order to quantify the relations between previous and posterior context we consider:

- the verbs related to the given adverbs
- the prepositional phrases selected by the temporal phrases
- the personal nouns and names in the context, according to [4]
- the punctuation signs, mainly commas

The overall results are presented in Table 1. The second column shows the number of examples obtained, after the elimination of phrases where there is no relation between the AdvT and the TimeN. Since the examples were automatically obtained from the snippet, some of them were not considered because of the lack of text when the sentences are split and context is omitted around the searched phrase. Column 4 gives the number of these eliminated examples because of short snippet. Columns 3 and 5 show the results after syntactic and semantic analysis of the context.

We manually analyzed the classified examples. Table 2 summarizes the results for the cases in Table 1 with more than 85% of age-related phrases. The column "Age-related" comprises the right and left context for the phrases denoting age of person; the percentage of phrases where context was identified and classified is shown in the first row. Column "No age" has the same structure. Table 3 summarizes the results for the cases in Table 1 with less than 85% of age-related phrases. The sequences of words and the part of speech appearing in the columns of right and left context correspond to the more general and more interesting patterns; for example, in the class *aún a sus NUM años* there are 34 phrases missing since they require many different patterns. The number of phrases is given in parentheses when it is bigger than one.

	Age-related		Not age		
	Left Right		Left	Right	
	context	context	context	context	
aún a sus NUM		88%		100%	
años		de vida (4)		de antigüedad	
		de existencia		de viajar	
		de edad (44)		de experiencia	
		cumplidos (4)		de tradición	
		VERB (148)		de viuda	
		NAMES (6)			
		PUNCT (40)			
aún hoy a sus	CONJ (23)	35%		100%	
NUM años	PUNCT (9)	de edad (4)		de partida	
		VERB (20)			
		PUNCT (11)			
aún a tus NUM		57%			
años		PUNCT (4)			
ahora a mis	61%	de edad (11)		100%	
NUM años	PUNCT (38)	PUNCT (39)		de casada	
	CONJ (72)				
aún con sus es-		100%			
casos NUM		PUNCT (4)			
años					
ahora de NUM	* N/P-PCT (233)	86%	71%	ahode list	
años	NAM/PN (18)	*PUNCT (233)	SER (27)		
		de edad (26)	periodo		
			disponer (4)		

Table 2. Results of context analysis for the classes with overwhelmingly age-related meaning

In the tables: VERB considers a verb related to the adverb, NAMES correspond to person's name, PUNCT comprises comma, parenthesis, semicolon, CONJ corresponds to conjunctions introducing new phrases, NAM/PN corresponds to name or personal noun, N/P-PCT comprises name or personal noun followed by elements of PUNCT, SER comprises "to be" conjugation, TENER and CONTAR, the "to have" verb. An asterisk (*) means that the left and right context is matched. {< 100} means NUM value lower than 100. {¡} means it excludes context for "No age" cases. Per-

cent marks in bold numbers indicate the quantity of examples in the context descriptions.

Table 3. Results for the classes with age-related meaning depending on context

	Age-related		Not age	
	Left context	Right context	Left context	Right context
aún con sus casi NUM años	PUNCT	100% VERB PUNCT (2) de edad		100% en el mer- cado de servicio de vigencia
aún con sus NUM años	**PUNCT (15)	89% **PUNCT (42) VERB (21) CONJ (3) <i>acsus</i> list		100% <i>No_acsus</i> list
de alrededor de NUM años	85% N/P-PCT (15) NAM/PN (117)	de edad (11)	51% <i>No_dealde</i> list	<i>dealde</i> list
actualmente con NUM años	N/P-PCT (62) NAM/PN CONTAR (27) NAM/PN que CONTAR (13)	48% de edad (98) de vida (2) recién cumplidos (3) cumplidos (2)		70% actcon list
actualmente de NUM años actualmente de	NAM/PN (5)	50% de edad (5) 100%	100%	
unos NUM años		de edad N/P-PCT (2)	SER (12)	
ahora con casi NUM años	; {CONJ, PUNCT} (30)	37% PUNCT VERB (21) VERB (34)	de abando- no(2)	74% {i} ahocca si list
ahora con más de NUM años		64% N/P PCT (8) VERB (16) de edad (4) encima	hasta (4)	100% <i>ahocmas</i>
ahora a NUM años		100% de edad		
ahora a los NUM años	porque (4)	46% de edad VERB (21)		18% ahoalos list
alrededor de NUM años	72% {< 100} TENER (28)	{< 100} de edad (4)	28% hace (47) durante (8)	
alrededor de los NUM años		17% de edad (51)		41% {< 100} (22

We can observe that the classes *aún a sus NUM años* and *ahora de NUM años* are the best examples for context identification. The worst case is *alrededor de los NUM años*, where we notice that almost all phrases indicate an age but in a general form, for example, *el consumo de frutas se da sobre todo alrededor de los 60 años* "the consumption of fruit is mostly seen about the 60 years old".

The class *ahora de NUM años* shows an interesting property: many age-related examples have right and left context matching that includes punctuation, isolating the temporal phrase and giving a context independent meaning.

Contexts for phrases not age-related share prepositional phrases modifying the noun time: for example, 7 *años de casada* "married for seven years", 7 *años de cárcel* "seven years in jail". In Table 3 two asterisks (**) means that for some examples the left and right context is matched

Some of the lists indicated in both tables by italics are enumerated in the following paragraphs and the number of cases is given in parentheses when it is bigger than one:

ahode: de antigüedad, de becas, de cárcel, de casados, de duración, más (5) *acsus*: encima (3), a cuestas (4), a las espaldas, cumplidos, de edad (14), de vida (5).

No_acsus. de amistad, de aplicación (2), de estudios, de existencia, de experiencia (3), de incapacidad, de profesión, de tradición, promedio (4)

dealde: de antiguedad (15), de duración (2), de evolución, de fallecido, de gobierno (2), de investigación, de investigar, de matrimonio, de persecución, de políticas, de prisión (2), de reformas, de trabajo, luz (9)

No_dealde: una antigüedad (7), datan (19), después (11), distancia (5), duración (2), SER (17), experiencia (3), lapso (2), luego (4), vida media (2), edad media, período (22), vida (3), vida conyugal

Such lists of phrases corresponding to local context presented in Tables 2 and 3 will be the basis for future work on lexical, morphologic, syntactic, and semantic categories for these temporal expressions' structural description. In this work we have used these contexts to automatically determine the temporal phrases reinforced by an adverb that correspond to age-related expressions in a sample taken from a second Mexican newspaper.

We search in this collection for the phrases: *aún a *años, ahora de * años, ahora con casi *años*, and *actualmente de *años*. We applied the results of context analysis in more than 300 examples and only 3 cases were not detected. The errors were due to false nominal phrase identification that corresponds to the person whose age is mentioned. Although the number of examples is low, we assume that the analyzed context is very similar for these kinds of phrases since the results are very near the 100 percent mark. Also, it should be noted that we used the Google search engine (www.google.com) for Spanish language so the analyzed context in Tables 2 and 3 corresponds to examples where several dialectal variations of Spanish are considered.

5. Possible Future Work

This paper has introduced a context-based method to automatically identify person's age phrases where adverbs make obligatory some conclusions and reinforce the mean-

ing of time in different forms. We suspected that these phrases were not confined to the Spanish language but that both language and culture influence their structures. Therefore a possible future work will consider applying the same method previously described to Italian and French languages where we found very similar examples. The following are some examples obtained from the Internet:

1. For French, we found examples similar to the Spanish phrases *alrededor de* NUM *años*, *aún a sus* NUM *años*, *aún a mis* NUM *años*, *ahora con* NUM *años*, *ahora a mis* NUM *años*, respectively:

Quand j''avais **autour de dix ans**, je me souviens qu'il y avait trois questions métaphysiques...

ça se gâte **encore à ses quinze ans** où il découvre qu'il est homosexuel.... ça promet...

En plus du travail, il y a le sourire, la décone, etc... J'adoooooreeee.... je serais chez toi encore à mes 95 ans...-)) Bisous

Maintenant à quatre-vingts ans, il faut stopper. — J'ai la moitié de fait. . . la meilleure moitié !

Il m'a dit qu'il ne voulait pas compromettre son image en emmenant son frère mineur en boîte, mais **maintenant avec mes 17 ans** c'est un peu comme si je ne ...

2. For Italian, we found examples similar to the Spanish phrases *aún a mis* NUM *años, ahora de* NUM *años, alrededor de los* NUM *años* :

Penso **ancora ai miei tredici anni e ai miei diciassette**, penso a quant'ero sana perché vivevo bene la mia sessualità, perchè scherzavo e ridevo come ...

Al punto che, ricorderò sempre una vacanza in Italia di cinque anni fa, quando scoprii che per mio figlio, **allora di otto anni**, una società omogenea era ...

Intorno ai vent'anni ho vinto una borsa di studio per frequentare l'Accademia di giovani autori di Mogol, è stato per me molto importante, ...

Although we suspect that other languages make use of the same syntactically relevant components, it is possible that they vary in terms of which component takes precedence over another and a manual intensive analysis should be performed. The analysis of other languages in the same way as in this paper will eventually give useful results for machine translation among such languages.

Another possible future work is research on speech language. Since phrases like aún con mis pocos NUM años, aún con mis cortos NUM años, aún con sus escasos NUM años, aún con tus casi NUM años, aún con sus casi NUM años are phrases which are used in everyday language and very few examples were obtained from the Internet, examples for these phrases and new examples should be obtained from oral corpora.

6. Conclusions

The variety in the structure of temporal expressions makes analysis of different combinations of classes of words necessary. We analyzed temporal expressions including the noun of time *year* that are modified by an adverb of time and the whole phrase expressing a person's age.

We first present a method to enrich the classes of temporal phrases when only a few examples are compiled. To obtain a more representative sample we compiled examples from the Internet for each class. We manually analyzed the context surrounding them to define the specific context for such expressions in order to automatically identify them. Specific context was obtained for nine of 21 classes.

The automatic identification of these phrases and their interpretation will directly benefit natural language processing tasks including: response to questions; visualization of events in lines of time; generation of phrases; translation, etc.

The temporal expressions considered present very interesting issues. The adverbs make some inferences obligatory and reinforce the meaning of time in different forms. Besides the fact of the time duration involved, they imply a direct judgment on the perception of the speaker, on the subject or both.

References

- Carbonero Cano, P.: Tiempo y aspecto en el adverbio (a propósito de "todavía", "aún" y "ya") Estudios lingüísticos en torno a la palabra. Universidad de Sevilla, Facultad de Filología (1993) 59–69
- Ferro, Lisa; Laurie Gerber, Inderjeet Mani, Beth Sundheim and George Wilson.: TIDES 2003 Standard for the Annotation of Temporal Expressions, MITRE Corporation (2004)
- Galicia-Haro, S.N.: Using Electronic Texts for an Annotated Corpus Building. In: 4th Mexican International Conference on Computer Science, ENC 2003, Mexico, (2003) 26–33
- Galicia-Haro, S. N., Alexander Gelbukh, Igor A. Bolshakov: Recognition of Named Entities in Spanish Texts. MICAI 2004. Proceedings in Lecture Notes in Computer Science (2972), Springer Verlag (2004) 420-429
- Gelbukh, A., Bolshakov, I.A.: Internet, a true friend of translator: the Google wildcard operator. International Journal of Translation (2006) 18(1–2), 41–48
- A. Gelbukh, G. Sidorov. Approach to construction of automatic morphological analysis systems for inflective languages with little effort. In: Computational Linguistics and Intelligent Text Processing, Lecture Notes in Computer Science, N 2588, Springer-Verlag. (2003) 215– 220
- 7. Kilgarriff, A.: Googleology is Bad Science. Computational Linguistics (2007) 33: 147-151
- Llido, D., Berlanga. R. y Aramburu, M.J.: Extracting temporal references to assign document event-time periods. Lecture Notes in Computer Science 2113 (2001) 62–71
- Mandel, Mark and Christopher Walter. Pautas para la anotación del Tiempo para Lenguas poco enseñadas (basado en los estándares de TIMEX2) Versión 1.0 Consorcio de Datos Lingüísticos (2006)
- Saquete, E., y Martinez-Barco, P. Grammar specification for the recognition of temporal expressions. In Proceedings of Machine Translation and multilingual applications in the new millennium. MT2000, Exeter, UK (2000) 21.1-21.7
- Saquete, Estela; Patricio Martinez-Barco, Rafael Muñoz. Recognizing and tagging temporal expressions in Spanish. In Workshop on Annotation Standards for Temporal Information in Natural Language, LREC 2002
- Wiebe, J. M.; T. P. O'Hara, T. Ohrstrom-Sandgren, and K. J. McKeever. An Empirical Approach to Temporal Reference Resolution. Journal of Artificial Intelligence Research 9 (1998) 247–293