In Proceedings of ICCM - 2007- Eighth International Conference on Cognitive Modeling. 7 - 12. Oxford, UK: Taylor & Francis/Psychology Press.

How Readers Retrieve Referents for Nouns in Real Time: A Memory-based Model of Context Effects on Referent Accessibility

Aryn Pyke (aryn.pyke@gmail.com)
Institute of Cognitive Science, Carleton University,1125 Colonel By Drive
Ottawa, ON K1S 5B6 Canada

Robert L. West (rlwest@connect.carleton.ca)
Institute of Cognitive Science, 1125 Colonel By Drive
Ottawa, ON K1S 5B6 Canada

Jo-Anne LeFevre (jlefevre@connect.carleton.ca)
Department of Psychology, 1125 Colonel By Drive
Ottawa, ON K1S 5B6 Canada

Abstract

When a reader encounters a noun, she tends to rapidly associate the noun with a mental referent (representation of entity in question). Our computational model confirms that a memory-based account is sufficient to account for a high rate of success at preliminary referent retrieval. Definite noun phrases ("The dog") can be used *anaphorically* to refer to referents already mentioned in the text, but they also frequently introduce a new referent into the mix (Poesio & Vieira,1998). An adequate model must account for how a reader makes an explicit or implicit decision about each noun's anaphoric status. We suggest that LTM contains both generic referent types and specific referent tokens, which simultaneously compete for retrieval via resonance. Our ACT-R simulation operationalizes the memory-based view to model the pre and post-noun activations of referents in memory. It predicts which referent will be retrieved (i.e. the most active), and consequently whether an anaphor will be initially treated as a new referent. The influence of anaphor word choice is explained, and encompasses metaphoric anaphors. Simulations results are congruent with human performance in our eye-tracked reading study, in which regressions to reanalyze an anaphor are indicative of the incidence of preliminary error.

Keywords: noun anaphora; memory-based text processing; resonance; reference assignment; cognitive modeling; ACT-R

Introduction

Evidence suggests that readers interpret incoming sentences incrementally, as the words become serially available (e.g., Sevidy, et al., 1999). In particular, human interpreters make a rapid preliminary association of a noun phrase (e.g., "the fruit) with a referent almost immediately after they encounter the noun (Just & Carpenter, 1980; see also Dell, McKoon, & Ratcliff, 1983, Sanford & Garrod, 1989). The term 'referent' is used here in the cognitive sense (as in Gundel, Hedberg, & Zacharski, 2001) to mean a mental representation of the person or object in question. At the time a noun is encountered, the processing system is not privy to information in the sentence that occurs after said noun. Thus, a reader's preliminary referent assignment is influenced by the preceding context and the noun itself. Our objective was to model this preliminary, on-line referent assignment for definite noun phrases. Such preliminary referent assignments may in turn be subject to subsequent adjustment. However, the present focus is on the nature and accuracy of the preliminary referent assignment process itself. The performance of the model will be compared with the accuracy of human data.

The Cognitive Task: Referent Assignment

To motivate the model, we first discuss the cognitive task (problem space) in more detail. In general, a noun can be used to introduce a new referent into the discourse (introductory use), or to refer to a referent that was previously discussed (anaphoric use). To enable us to process these two uses, we have two types of referents in memory: representations of specific people and objects that the cognitive agent is already familiar with (e.g., the particular apple in your bag), and more generic/prototypical representations (e.g. a generic apple). The latter, generic referents are appropriate to retrieve when the noun is used in an introductory capacity. For example, in (2) "fruit" is used anaphorically to denote the referent introduced by the antecedent "apple" in (1). Note that an anaphoric noun need not match the label that was previously used for that referent, and the same noun "apple" can be used to refer to different referents in (1) and (3).

(1) John bought an apple.
(2) John bit into the fruit.
(3) The apple Mary bought was green.

A more fundamental issue is that a definite noun, "The apple" bears no explicit indication of whether it used in an introductory or anaphoric capacity. The definite article "the" was often assumed to indicate that the intended referent is already familiar to the reader (e.g., Clark & Sengul, 1979; Garnham, 1989; Just & Carpenter, 1987). However, as exemplified in (3), corpus analyses have established that definite noun phrases are equally likely to introduce new referents into the discourse as to denote referents that have already been mentioned (Gundel et al., 2001; Poesio & Vieira, 1998). Thus, readers are not privy to the anaphoric status of a definite noun a priori. Consequently, during preliminary assignment, readers may sometimes misinterpret an anaphoric noun as if it is introducing a new referent (or an introductory noun as anaphoric). In the literature, the reader's task upon encountering a definite noun phrase is sometimes dubbed *anaphor resolution*, but this terminology clearly under-represent the full scope of the reader's (and modeler's) task. Thus, the job of the present model is to describe the means in which a referent for a noun (in a given context) is rapidly selected on-encounter. And in so doing, the model should serve to predict the likelihood that the selected referent will be the correct one (i.e., will it be generic-new or specific).

The present treatment applies to the preliminary assignment of referents to definite *nouns*. As Gernsbacher (1989) suggests, that this memory-based process may apply to other types of referring expressions like names (John) and pronouns, (he). However, this issue goes beyond our present data. Furthermore, in contrast to nouns, which can be introductory or anaphoric, pronouns are almost exclusively anaphoric (Kintsch, 1998), and are subject to more syntactic constraints, so there is reason to believe that the problem space and process may be somewhat different. Consequently, appropriate referents for non-noun referring expressions are effectively hard-coded in our simulation.

Criteria for a Cognitive Referent Retrieval Model

Reference assignment is addressed by some non-cognitive Natural Language Processing algorithms (e.g., Bean & Riloff, 1999; Vieira & Poesio, 2000), however, they involve multiple passes forward and back through the text, and do not directly speak to the development of a *cognitive* model of *on-line* processing. Thus, to set the stage for the proposed model, we layout the general criteria for a psychologically plausible model of referent retrieval for noun phrases.

1. **Incremental Processing:** The system should be able to make use of information in each incoming sentence incrementally, roughly word-by-word.
2. **Appropriate Representational Units**: In line criteria 1, the unit of analysis for the processing system must be smaller than a complete sentence. The cognitive task being modeled is to associate a definite noun phrase with a mental representation of its referent, so the system must, minimally, have representations (though possibly atomic ones) for individual nouns and potential referents, in order to model the task of associating the former to the latter.
3. **Context Sensitivity:** The model should allow and account for the influence of: (i) preceding context sentences, and (ii) preceding parts of the current sentence, and (iii) the current noun itself, on the processing of the current noun. To account for the processing of a particular anaphor, the model should also, as a pre-requisite, model the processing of the prior context, to get the memory system into the appropriate state (so relative accessibilities of specific and generic referents reflect the influence of prior context).
4. **Real-Time Simulation**: Ideally, the reading process should be simulated in real-time units (e.g., ms vs. 'cycles'). Cognitive effects like priming (spreading activation) are time-sensitive and subject to decay. The system should ideally model such memory effects (fluctuations in activation over time), and should simulate processing of the text at a representative reading rate (e.g., 150 ms/word).
5. **Appropriate Problem Space**: Models necessarily abstract away from some level of detail, but care must be taken to ensure that the problem space in the model is not artificially skewed or trivially sparse. The characterization of the task and the representation of the problem space (e.g., range of possible referent choices) should be sufficiently rich to reflect the interpretation challenge facing a real reader, and so permit key types of possible errors. Since readers do not know a priori whether a current definite noun phrase is anaphoric or introductory, the explicit or implicit determination of this property is part of the referent assignment process. Thus, the model should ideally be able to operate on and discriminate between (though not always accurately) both anaphoric and non-anaphoric definite noun phrases. In this vein, the memory system must be populated with not only the correct referent, but also the other referents in the discourse, and generic referent prototypes.

We are unaware of a previous model that fulfills all of these criteria. Other models (e.g. Budiu & Anderson, 2004, Lemaire & Bianco, 2003) fulfill some but not others, notably criteria 2 and 5. In view of these criteria, the model was implemented in ACT-R (discussed later). Next, we outline the theoretical framework underlying our model.

The Framework: Memory-based Processing

Our model is inspired by the memory-based view of text-processing (see Gerrig & O'Brien, 2005 for a review). In particular, we propose that preliminary referent retrieval is driven by general-purpose memory mechanisms. In particular, under the resonance model (e.g., Gernsbacher, 1989; Myers & O'Brien, 1998) current information in working memory (i.e., the anaphoric noun) serves as a cue that automatically boosts activation of other entities throughout long-term memory -- including, ideally, the intended referent -- in accord with their conceptual overlap with the cue. Thus, at the time the anaphor 'fruit' in (2) is encountered, the apple referent can be automatically re-activated via resonance in virtue of its conceptual overlap with the anaphor (a pre-existing conceptual association). In our model (see also see also Budiu & Anderson, 2004), the strengths of conceptual associations were estimated using Latent Semantic Analysis (LSA) values (lsa.colorado.edu; see Landauer, Foltz, & Laham, 1998 for a review), which give a maximum similarity of 1 (i.e., similarity of a concept to itself). So had the anaphor used been "apple" it would have provided an even larger activation boost to the intended referent [R1:apple], than was provided by the anaphor "fruit" (LSAs:<apple,fruit>=.47,<apple,apple> =1). Thus choice of noun itself exerts an immediate impact on the relative accessibilities of referents via resonance.

This memory-based account can be contrasted with a more special-purpose process of referent retrieval. Some suggest that when a reader encounters an anaphoric noun, he/she undertakes a proactive search for a referent mentioned within the current text (e.g., Clark & Sengul, 1979; Kintsch & Vandijk, 1978; O'Brien, Plewes, & Albrecht, 1990). The discourse might be mentally represented as a series or network of propositions, and the reader might systematically troll backwards through it in search of a referent that (according to some criterion) could constitute a match to the current anaphor term. However, since many definite nouns not anaphors such a proactive, process-of-elimination search would often be a waste of time. Further, if general memory mechanisms are often sufficient to automatically bring a referent to mind, parsimony argues against the proposal of a proactive special-purpose search process. The model in this paper confirms that memory-based accounts are sufficient to account for a high rate of success at preliminary referent retrieval for anaphoric nouns. The theoretical and empirical arguments against a special-purpose search account are outlined more fully in Pyke, West and LeFevre (2007).

In our model, generic referents and specific discourse referents simultaneously compete for retrieval. The most active referent is retrieved. Thus,it is *not* the failure to find/retrieve a referent that then, serially, leads to treating a noun as a new referent. Rather, retrieval, per se, typically always succeeds. Whether it is a specific or generic referent that is retrieved determines whether the noun as treated as anaphoric or introductory in preliminary analysis.

Factors Affecting the Activation Levels of Referents

Each referent's accessibility during preliminary noun processing owes to two components: i) the activation boost/spread from the noun term currently being processed; and ii) its 'context dependent' pre-noun activation level. Such context factors are outlined below.

1.**Spread of Activation from Pre-Anaphor Words**. Just as the anaphor resonates with, or spreads activation to, referents, our model assumes that such activation spread generally occurs as each content word in the discourse is encountered. The activation boosts received by referents may persist (as do lexical priming effects, e.g., Collins & Loftus, 1975) even when the reader progresses on to the next word. While such effects decay they may exert a cumulative effect on a referent's activation.

3.**Recency and Frequency of Use of the Referent**. These general factors affect any mental representation's accessibility. Evidence indicates that the further back that an antecedent is (in sentences, and consequently in time), the more challenging it is to process the anaphor (e.g., Clark et al., 1979; Duffy & Rayner, 1990; Levine, Guzman, & Klin, 2000). A referent referred to many times in a text, and/or referred to in the sentence preceding the critical anaphor should be more active, ceteris paribus, than a referent mentioned only once several sentences back.

4.**Sentence Wrap-Up Effects**. Just and Carpenter (1980) suggested that integrative processes occur at sentence end, which is why readers spend tend to spend relatively longer on the final word in each sentence. These wrap-up processes may result in sentence-end activation effects (Balogh, Zurif, Prather, Swinney, & Finkel, 1998). Probe studies suggest that a referent mentioned early in a sentence may also produce facilitation effects at sentence end (e.g., Dell et al., 1983; McKoon & Ratcliff, 1980). In our model, the processes at sentence end result in an activation boost of the specific referents mentioned in the sentence.

5. **Discourse Dependent Associations**. In addition to pre-existing associations like those we are modeling with LSA, discourse dependent associations may be formed in memory. Spread of activation through such associations may produce intermittent (yet cumulatively significant) activation contributions to an intended referent during pre-anaphor processing. For example, each sentence (and proposition) in a discourse may contain several referents. In Dell et al. (1983, see also McKoon & Ratcliff, 1980), two referents which have appeared in a common sentence are called *companions*. The comprehension process appears to forge a lasting association between companion referents in memory, possibly during sentence wrap-up, such that when a referent is subsequently encountered, its companion(s) from prior sentences also become re-activated right away, and to a comparable degree (Dell et al., 1983). Thus, in our model, whenever a referent is mentioned, its companions are also boosted in activation, thereby making them more accessible as referent candidates for up-coming nouns.

Our treatment extends the prior treatments in the literature in that it provides a more explicit, comprehensive, quantitative and real-time operationalization of such contextual influences on the LTM referents' pre-anaphor activation levels. Furthermore, most discussions of noun anaphor processing in the literature (c.f., Garrod, Freudenthal & Boyle, 1994, Levine et al, 2000), fail to address the fact that the reader does not know that a noun is an anaphor a priori (The classification problem, see also Pyke, West & LeFevre, 2007), so the incidence of new-referent errors during preliminary assignment has been largely unexplored and, in our view, underestimated.

The Data: Human Performance

How likely are readers to associate an anaphoric noun to a new referent on encounter? It was originally believed that people made almost no preliminary errors in assigning a referent to noun anaphors (Sanford & Garrod, 1989), however subsequent evidence has established that such errors do indeed occur (e.g., Levine, Guzman & Klin, 2000, and the present research). Such errors were elicited by manipulating factors related to the accessibility of the intended referent (e.g., the referent had not been mentioned for several sentences).

Because a reader's preliminary referent assignment for a noun occurs on-line, (mid-sentence), and is not directly observable by an experimenter, information about the nature (accuracy) of such preliminary assignments is somewhat challenging to empirically obtain. Note that the reader's final interpretation at sentence end may reflect the influence of subsequent processes and information. We observe that a reader's eye-movements can provide a useful indication of the effectiveness of preliminary referent assignment. In the course of processing the remainder of the sentence after the anaphor, errors in preliminary assignment are often detected. In general, when readers detect that they have made earlier errors in interpretation, they often regress their eyes back to the site of the initial misinterpretation to do an *overt reanalysis* (e.g., Altmann, Garnham, & Dennis, 1992; Meseguer, Carreiras, & Clifton, 2002). We conducted an eye-tracked reading study to determine the relative likelihood of regressions to reanalyze the anaphor, when we manipulated the choice of anaphoric noun.

Experiment

The stimuli were 42 stories that were each 4 sentences long. The fourth, critical sentence commenced with a definite noun phrase: "The (noun)", and this noun was intended anaphoricallly to denote a target referent introduced in the first or second sentence. For example:

> Mary had a pet terrier.
> It was white and shaggy.
> She took it to the beach.
> The terrier/dog/mop barked at the birds.

Anaphor word choice was manipulated so each of the 42 stories had three different versions: i) antecedent-match (terrier-terrier), – anaphor was identical to the noun which originally introduced the referent ii) category (terrier-dog); and iii) metaphoric (terrier-mop). Recent research suggests that the same general mechanisms apply to the processing of literal and figurative content (Budiu & Anderson, 2004, Giora, 2002; Glucksberg, 2003; Kintsch, 1998, chap. 5.3). The memory-based model is compatible with this claim. Activation spreads automatically in accord with similarity, so just as activation spreads from the anaphor dog to the referent [R1:white, shaggy, terrier], activation should also spread, from the anaphor mop to the mental representation of the white, shaggy terrier. The latter case provides less spread of activation (LSAs: <dog, white-shaggy-terrier>= .18, <mop, white-shaggy-terrier>=.07), so readers should be more prone to make a preliminary new-referent assignment errors for the metaphoric anaphors.

Procedure. Stories were presented line-by-line to participants (N=24), while an EyeGaze™ System tracked their right eye. Story versions were counterbalanced, so each participant saw only one version of each story, and thus saw 14 stories of each anaphor type.

Results. Eye-movements were analyzed for the critical sentence. First-pass reading time for the anaphor did not vary for the different noun versions (p>0.05, word length and frequency as co-variates), however, the likelihood of later regressing the eye back to the anaphor differed, $F(3,71.9) = 7.51$, $MSE = 1.60$, $p = .000$. Readers made a regression back to the anaphor on 52% of metaphoric anaphors, 36% of category anaphors and 41% of antecedent-match anaphors (see also Figure 2).

Discussion: Regressions to the antecedent-match and the category anaphors were comparable, though, somewhat surprisingly, more regressions were made to the antecedent-match anaphors. Readers regressed most to the metaphorically intended anaphors, presumably because they had made many preliminary assignment errors and treated them as new referents. This explanation was confirmed in a follow-up cloze study. Readers were presented with the first 3 sentences and the critical noun phrase (e.g., The mop_____________), and created their own completion of the sentence. The completions revealed whether the reader had associated the noun with the intended referent or had treated it as a new referent. In 48% of the trials, readers associated the metaphoric anaphor with a new referent.

The Model

In the model, the referent retrieval process in play while reading the anaphoric noun is basic, blind and strictly memory-based: When 'reading' the noun, activation automatically spreads from the noun to all referents in memory (both generic and specific), and then the most active referent from memory is retrieved. The impact of prior discourse processing will be entirely mediated by its lingering effect on the activations of the various referents in LTM. Thus, to model the preliminary referent assignment for a particular noun, we must also simulate the processing of the preceding discourse, but only so far as is necessary to approximate its effect on the immediately pre-anaphor activation levels of the various referents in LTM.

The 3 versions of each story are identical up until the anaphor. Thus the pre-anaphor activation levels of the intended referent and other specific referents in the story are the same across versions (say, [R1:Mary]: 2.4, [R2:terrier]:1.4, [R3:beach]:2.8). What about generic referents? The anaphor term determines the relevant (most competitive) generic referent in play. If the anaphor is "mop", the intended referent [R2:terrier] competes not just with the other story referents but also with the generic referent [G1::mop]. For the version of the story with the anaphor dog, the relevant generic competitor is [G2:dog]. The activation levels of the generic competitor will depend on the spread of activation it receives from the pre-anaphor words and the anaphor itself. In contrast, specific referents in the story also get activation boosts from sentence-wrap up and companion spreading.

Overview of Operation

0. LTM is seeded with generic referents for various discourse concepts, including, importantly the antecedent concept and also (if different) the anaphor concept. For the "mop" version of our example story, memory will be seed with [R1:Mary], [G1:terrier], [G2:Beach], [G3:birds] and [G1:mop]. If, upon reading

1. Words of a story are then processed serially.

2. Each content word (e.g., noun, adjective, verb) automatically spreads activation to both specific and generic referents in memory according to the LSA similarity between the word and the referent.

3. If the current word is a referring term (noun, name, pronoun), a referent is retrieved from memory. The referent retrieved will be the most active one, be it specific or generic. In the latter case, the generic referent is used to create a new specific referent to associate with the noun.

The assigned referent is automatically boosted in activation (in virtue of it's current use), and activation also spreads from it to its companion referents from previous sentences. A fan-effect applied, so if the current referent had n companions in the previous sentence, the weight factor for activation spread to each companion will be (1/n).

4. At the end of a sentence during wrap-up, the sentence's referents become reactivated and mutually associated.

Besides the original descriptor used to introduce a referent (e.g., "terrier"), other properties can also be explicitly mentioned in a text (e.g., white & shaggy). In step 2 above, activation from incoming words spreads to each *explicit* attribute of each referent [R1:terrier,white,shaggy]. Each attribute then spreads activation to the referent representation as a whole. The more active (recently mentioned or primed) an attribute is, the stronger its relative contribution (weight = attribute's activation level * LSA<current word, attribute>).

Implementation Architecture: ACT-R

The criteria outlined previously motivated the choice of ACT-R (Anderson et al., 2004) as a suitable cognitive modeling platform for our model. In particular, we used the python extension of ACT-R (Stewart & West, in press).

The ACT-R architecture is predicated on a Unified Theory of Cognition (Newell, 1987) - that is, on the belief that our performance on a vast range of tasks can be accounted for parsimoniously by a common cognitive system operating with general-purpose mechanisms and principles. As such, ACT-R is spiritually compatible with the memory-based framework in which referent retrieval is attributed to general-purpose memory mechanisms.

Computer implementations do not inherently impose any psychological constraints on the character of the model, nor do they necessarily simulate the process in real-time. Consequently, the cognitive plausibility of one-off task-specific models is sometimes open to question. ACT-R has in-built psychologically motivated constraints, though it does have some 'arbitrarily' adjustable parameters. However, the architecture has proved conducive to modeling a vast range of cognitive tasks, and the data accrued has provided theoretically and empirically motivated constraints for the values/ranges for its key parameters. In our model, key parameters are set to recommended 'universal' defaults (e.g., noise=0.3, decay=0.5, production time=50 ms).

ACT-R supports two concurrent levels of functionality:

(i) a production system that carries out sequences of situation-driven productions (i.e. if-then rules) that serve as the procedural building blocks (steps) for various tasks.

(ii) a dynamic memory system which contains the various respresentational units (called chunks) upon which the productions act. The memory system can simulate the real-time fluctuation of activation of a representational unit. For example, the effect of recency and frequency of use on the activation level of representation R_i is quantified by B_{Ri}, where tj is the time since use j, and d is a constant whose default value is .5 (Anderson et al., 2004).

$$B_{Ri} = \ln \sum_{j=1}^{N} \text{tj\^{}-d}$$

Thus, ACT-R affords sufficient functionality to fulfill almost all of the desired criteria: i) incremental processing: cognitive operations (productions) can be executed at a rate of one per 50ms, so a model can perform several operations (e.g. recognize word, spread activation, retrieve referent) during the typical reading time for each incoming word (150 ms); ii) representational units – the modeler can specify the types and numbers of representations in memory, for example, noun-chunks, generic referent chunks, and specific referent chunks, iv) problem space – the architecture allows the modeler to define the task (set of productions) and populate memory (set of chunks) as appropriate, iii) real-time simulation – the rate of productions is paced to reflect the time for the mind to perform a single simple operation, and the memory system can simulate the fluctuations in activation of the chunks (referents) over time.

The content that is currently in the system's focus (e.g., the noun-chunk), can spread activation to other chunks (e.g., referent-chunks) in memory. However, in traditional ACT-R the boost in activation received by a chunk (referent) from a stimulus (word) is removed as soon as the stimulus word is no longer in focus. This precludes any priming effects of spreading activation from prior words. To allow for such persistent effects in our model, we introduced this functionality into python ACT-R. In so doing, we feel we've augmented rather than circumvented the psychological plausibility of the architecture.

Simulation Results

The model was run 100 times on each version (match, category, metaphor) of each of 42 stimuli stories. Figure 1 depicts mean pre and post anaphor activation levels of the intended referent (R) and the relevant generic competitor (G). Match anaphors produced the highest, absolute post-spread R-activations. The simulation revealed considerable variation in pre-anaphor accessibility of the target referent from story to story. However, the pre-anaphor R-activation is the same for all 3 versions of a given story. In general, the greater the activation ‘head-start’ a specific intended referent has due to pre-anaphor context influences, the greater the latitude in anaphor word choice.

Figure 2 indicates how frequently the simulation retrieved the correct referent. These results are correlated with the likelihood of regression for the 42*3 items in the human data ($r=-.333$, $p=.000$). For match anaphors, humans regressed on more trials than would be expected in light of the minimal number of preliminary referent retrieval errors predicted by the model (<10%). Regressions in the antecedent-match case may result not only from preliminary errors but may be inflated due to a pragmatic "repeated name effect" (Kennison & Gordon, 1997). When a reader does regress, they may be re-engaging the memory-based referent retrieval process, at a time when the referent's accessibility may be boosted due to priming from post-anaphor words. We will extend our simulation to test this.

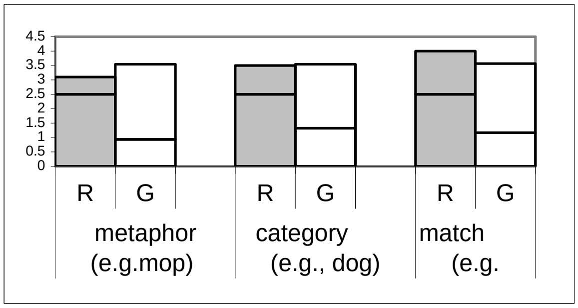

Figure 1: Pre(lower half) & Post-spread activation level for the intended referent (R) and its generic competitor (G).

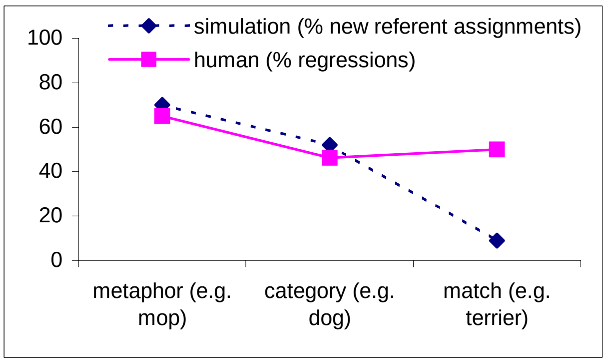

Figure 2:Comparison of Simulation & Human Performance

Concluding Remarks

Our model operationalizes the memory-based view to estimate the (pre & post-noun) accessibilities of referents in memory, and thus predicts when a particular anaphor will be initially misinterpreted as a new referent. Such a simulation tool could have a practical application to assess how comprehensible (each referring expression in a) text is. And in psycholinguistics research to check whether accessibility levels are controlled/comparable across stimuli. Although readers can often later correct preliminary errors, their final representation of the text may be degraded since vestiges of misinterpretations persist in memory (Johnson & Seifert, 1998). Further research is necessary to explore such potential long-term 'costs' of using anaphors that are difficult to resolve during preliminary analysis.

Acknowledgments

We express our thanks to Terry Stewart and NSERC.

References

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. L. (2004). An integrated theory of the mind. *Psychological Review, 111,* 1036-1060.

Balogh, J., Zurif, E., Prather, P., Swinney, D., & Finkel, L. (1998). Gap-filling and end-of-sentence effects in real-time language processing. *Brain and Language, 61,* 169-182.

Bean, D. L. & Riloff, E. (1999). Corpus-based identification of non-anaphoric noun phrases. *In Proceedings of the 37th Annual Meeting of the Assoc. for Computational Linguistics.*, 373-380.

Budiu, R. & Anderson, J. R. (2004). Interpretation-based processing: a unified theory of semantic sentence comprehension. *Cognitive Science, 28,* 1-44.

Clark, H. H. & Sengul, C. J. (1979). In search of referents for nouns and pronouns. *Memory & Cognition, 7,* 35-41.

Dell, G., McKoon, G., & Ratcliff, R. (1983). The activation of antecedent information during the processing of anaphoric reference..*J. of Verb. Learning and Verb. Behav 22*,121-132

Duffy, S. A. & Rayner, K. (1990). Eye-movements and anaphor resolution - Effects of antecedent typicality and distance. *Language and Speech, 33,* 103-119.

Garrod, S., Freudenthal, D., & Boyle, E. (1994). The role of different types of anaphor in the online resolution of sentences in a discourse. *Journal of Memory and Language,* 33, 39-68.

Gerrig, R., & O'Brien, E. (2005). The scope of memory-based text processing. *Discourse Processes,* 39(2&3), 225-242.

Gernsbacher, M. A. (1989). Mechanisms that improve referential access. *Cognition, 32,* 99-156.

Gundel, J., Hedberg, N., & Zacharski, R. (2001). Definite descriptions and cognitive status in English. *Journal of English Language and Linguistics, 5,* 273-295.

Just, M. A. & Carpenter, P. A. (1987). *The psychology of reading and language comprehension.* Boston: Allyn and Bacon, Inc.

Kennison, S. & Gordon, P. (1997). Comprehending Referential Expressions During Reading. *Discourse Processes, 24,* 229-252

Kintsch, W. (1998). *Comprehension: A Paradigm for Cognition.* New York: Cambridge University Press

Kintsch, W. & Vandijk, T. A. (1978). Toward a model of text comprehension. *Psychological Review, 85,* 363-394.

Landauer, T.,Foltz, P.& Laham,D.(1998).An introduction to latent semantic analysis. *Discourse Processes,25*, 259-284

Lemaire, B. & Bianco, M. (2003). Context effects on metaphor comprehension: Experiment and simulation. In *Proceedings of 5th Int.Conf. on Cognitive Modelling (ICCM),* 153-158

Levine, W. H., Guzman, A. E., & Klin, C. M. (2000). When anaphor resolution fails. *J. of Mem. and Language, 43,* 594-617.

Myers, J. L. & O'Brien, E. J. (1998). Accessing the discourse representation during reading. *Discourse Processes,26,* 131-157.

O'Brien, E. J., Plewes, P. S., & Albrecht, J. E. (1990). Antecedent retrieval-processes. *JEP: LMC, 16,* 241-249.

Poesio, M. & Vieira, R. (1998). A corpus-based investigation of definite description use. *Computational Linguistics, 24,*183-216.

Pyke, A., West, R. L., & LeFevre, J. (2007). On-line reference assignment for anaphoric and non-anaphoric nouns. In *Proc. of* the 29th Annual Meeting of the Cognitive Science Society.

Sanford, A. J. & Garrod, S. C. (1989). What, when, and how?: Questions of immediacy in anaphoric reference resolution. Language and Cognitive Processes, 4, 235-262.

Sedivy, J., Tanenhaus, M., Chambers, C., & Carlson, G. (1999). Achieving incremental semantic interpretation through contextual representation. *Cognition, 71,* 109-147.

Stewart, T.C., West, R.L. (in press). Deconstructing and Reconstructing ACT-R: Exploring the Architectural Space. *Cognitive Systems Research.*