A naïve, salience-based method for speaker identification in fiction books

Kevin Glass and Shaun Bangay

Department of Computer Science
Rhodes University
Grahamstown

[k.glass;s.bangay]@ru.ac.za

Abstract

This paper presents a salience-based technique for the annotation of directly quoted speech from fiction text. In particular, this paper determines to what extent a naïve (without the use of complex machine learning or knowledge-based techniques) scoring technique can be used for the identification of the speaker of speech quotes. The presented technique makes use of a scoring technique, similar to that commonly found in knowledge-poor anaphora resolution research, as well as a set of hand-coded rules for the final identification of the speaker of each quote in the text. Speaker identification is shown to be achieved using three tasks: the identification of a speech-verb associated with a quote with a recall of 94.41%; the identification of the actor associated with a quote with a recall of 88.22%; and the selection of a speaker with an accuracy of 79.40%.

1. Introduction

1.1. Problem Statement

Dramatic scripts can be automatically generated from fictional books by extracting portions of text representing speech, provided that the identity of the corresponding speaker can be established. Speaker identification requires that the speech verb and its associated noun be discovered in order to determine the identity of the speaker. This paper investigates a novel salience-based technique for the identification of the speaker of quotes found in fiction books, without the use of complex machine learning or logic-based techniques.

1.2. Background

The process described in this paper is a component of a Text-to-Scene conversion system, where information from natural language texts is used to populate three-dimensional virtual worlds. Speech articulated by avatars in a fiction book is converted to audio, using a unique voice for each avatar that is speaking. However, the speaker of a quote is not always consistently nor explicitly indicated for instances of direct speech. The speaker may be referred to indirectly, using a pronoun or descriptive noun. Sequences of unattributed quotes may also exist that depend purely on the context of the quote to identify the correct speaker.

A number of clues exist within the surrounding text of a quote that aid in identifying the correct speaker. The first clue is the main verb that indicates speech, or the *Speech-Verb*. If this verb is correctly identified then one of the arguments of this verb, namely the *Actor* of the verb, points (either directly, or by means of anaphora) to the correct *Speaker*. The Actor differs from the Speaker in that the Actor is a token in the text which is the argument of a verb, while the Speaker is an avatar which is chosen from a list of avatars participating in the story. The Actor may be an alias of a Speaker (for example the Speaker's first name), but may also be an indirect reference in the form of a common noun ("the boy"), or pronoun ("he"). In cases where no verb exists, then the larger context of the quote must identify the Speaker.

The process of extracting quotes and associated speakers, henceforth referred to as *script extraction,* consists of the following phases:

1. Locating the Speech-Verb for the quote;
2. Locating the argument of the Speech-Verb that points to Speaker (referred to as the Actor);
3. Selecting the Speaker from the list of avatars that are participating in the scene.

Fiction books describe people, objects and actions that occur in a fictional or *virtual* world. The virtual world is filled with *entities* including avatars and objects. Fiction text refers to these entities using textual *aliases.* Every time an alias for a particular avatar is encountered in the text, the corresponding avatar in the virtual world must be identified. Different terms or portions of the text may refer to a single entity. This research assumes the existence of a previously constructed avatar list (created manually or automatically using *named entity extraction* processes [1, 2]). The gender of each avatar in the list is indicated, and each avatar is associated with a (non-comprehensive) list of aliases.

1.3. Overview

This paper is structured as follows: script extraction, and the inspiration behind the salience-based technique is described in Section 2. A description of the features used for identifying Speech-Verbs, Actors and Speakers is presented in Section 3, with the results of each phase presented in Section 4.

2. Related Work

This paper serves as an extension of previous work by Glass and Bangay [3] which presents a rule-based learning system for accomplishing the script extraction task. Although the rule-based system accomplishes the task successfully (achieving recall over unseen text of up to 95.17% for locating the Speech-Verb, 88.18% for locating the Actor, and accuracy of up to 82.85% for selecting the Speaker), it is dependent on the existence of annotated data for training. This work differs in that no training is required, and uses only a combination of commonly available natural language processing tools.

Script extraction from fiction texts is previously examined by Zhang *et al.* [4], where the Speaker of each quote is achieved with accuracies of between 47.6% and 86.7%. A number of issues are not discussed, for instance how different aliases are identified as the same Speaker, and how sequences of quotes with no explicitly identified Speakers are handled.

The salience-based method is inspired predominantly by work done in the pronominal anaphora resolution domain [5, 6, 7], which is concerned with determining the *antecedent* or referent of each pronoun in natural language. In anaphora resolution, salience-based methods function by locating tokens that are potential candidates for being the antecedent of a pronoun, and calculating a score for each candidate token according to certain features, including part-of-speech and other similar indicators. The token with the highest score is selected. A similar process is employed for script extraction, using salience-based methods for locating Speech-Verbs and Actors.

In addition to salience-based techniques, a number of hand coded rules are used for selecting the Speaker of each quote, based on the features of the Speech-Verbs and Actors identified. This is also inspired by anaphora resolution literature [8].

3. Clues for Speaker identification

The starting point for the script extraction process involves identifying quoted speech in the fiction text. These can be readily identified using the surrounding punctuation. For every quote identified, the Speech-Verb must be found, and thereafter the Actor. The Actor is then used to locate the Speaker of the quote. Parts-of-speech information is provided for each word in the input text, identified by an ensemble of parts-of-speech taggers [9], as well as parse information provided by the Connexor FDG parser [10].

3.1. Speech-Verb Annotation

To locate the Speech-Verb for each quote, the surrounding sentences of each quote are scanned for verbs. This results in a list of candidate tokens that may all possibly act as the Speech-Verb for the quote under consideration. A number of *features* are used to identify which candidate verb is the Speech-Verb. Each feature is assigned a salience value, and if a candidate exhibits a feature, then the salience of that feature is added to the total score of the candidate. The candidate with the highest score is chosen as the Speech-Verb. The features used are as follows:

- **Main verb:** salience is immediately awarded to tokens that are marked as main verbs by the FDG parser.
- **Hypernym:** a similar method as used by Velardi *et al.* [11] is employed, where salience is awarded to verbs that have `communicate`, `verbalise` or `breathe` as ancestors in a hierarchical lexical tree. WordNet [12] provides this information. All senses of the verb and their corresponding hypernym trees are scanned for one of these three ancestors. `Breathe` is also chosen as one of the ancestors since verbs such as `sigh` and `gasp` are often used to indicate speech, even though they have no lexical link with `communication`.
- **Adjacent Sentence:** verbs found within the adjacent sentence are awarded extra salience.
- **Proximity to quote:** verbs that occur in the preceding and following sentences are given extra salience based on their proximity to the quote. That is, words closer to the quotation receive a higher salience than words further away.

All features are awarded 1 point, with the exception of the proximity feature, which adds an amount proportional to the distance of the verb from the quote as a value between 0 and 1.

She **shuddered** when she **heard** little Jammes speak of the ghost , **called** her a " silly little fool " and then , as she was the first to believe in ghosts in general , and the Opera ghost in particular , at once **asked** for details :

" Have you seen him ? "

Figure 1: Example Speech-Verb, from *The Phantom of the Opera* by Gaston LeRoux

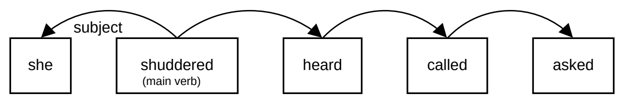

Figure 2: Visualisation of verb chain produced by the FDG Parser

3.2. Actor Annotation

Once the Speech-Verb for each quote is located, the argument of the identified verb which indicates the Actor needs to be identified. The Actor identification process is based on the idea that there exists a dependency between the Speech-Verb and the Actor that can be identified by the parser. In some instances *verb-chains* occur where a number of verbs are linked to one single main verb, that has as its subject or object the Actor of the Speech-Verb. The sentence presented in Figure 1 is an illustration of one such occurrence, with `asked` as the Speech-Verb. Figure 2 presents the verb-chain from the word `asked`, as derived from the dependencies indicated by the FDG parser. The chain extends to the main verb of the sentence from which the actor `She` can be derived.

The Actor annotation algorithm traverses up the verb-chain until the root is found. Thereafter the children of the root are examined for instances containing a *subject* or *object* relation, as identified by the parser. No text occurring inside quotes is considered.

Errors in the parsing of complex sentences often cause incorrect dependency structures from the FDG parser. As a result, a salience-based technique is used to locate the Actor, of which the verb-dependency is just one feature that is rewarded. This allows for the Actor to be located where no such dependencies are identified by the parser, or for the correct Actor to be chosen in cases where the dependencies are incorrect. The process begins at the Speech-Verb token. From this point tokens are traversed in both directions from the verb, and assigned ranks based on a number of features:

- **Subject or Object:** the token is awarded salience for being the subject or object of the main verb (at the top of the verb chain).
- **Parts-of-speech filter (noun/pronoun):** the token is awarded salience if it is a noun that is a descendant of the word: `person` (using WordNet [12]) or is not a recognised *English* noun (using the 12dicts English gazetteer [13]) which means that the token may be a name. In addition, a point is awarded for pronouns.

- **Proper Noun filter:** the token is awarded salience for being capitalised, unless it is the first token of a sentence. If a token is the first in a sentence, then it must not be recognised as English to be rewarded. Words that are marked as prepositional-complements are not awarded salience (for example *Joe*, in *He said to Joe...*).
- **Abbreviation filter:** often abbreviations such as Mr. or Mrs. are selected as candidate Actors. Such abbreviations are identified using an abbreviation gazetteer, and not awarded salience.
- **Distance from verb:** salience is awarded to those candidates which are closer to the speech verb.

All features are awarded 1 point. The exception is the distance feature, which increments the score of the candidate by multiplying the existing score by 0.1% for each word between the verb and the Actor token.

3.3. Speaker Resolution

The Actor token identified in the previous section is used to determine which avatar in the avatar list is responsible for the speech. However, the Actor token may occur in different forms:

- **Direct reference:** where the token directly identifies the Speaker, for example, *John* in *John said, "..."*.
- **Pronominal anaphora:** where the speaker is indirectly referenced using a pronoun, for example, *he* in *he said, "..."*.
- **Nominal anaphora:** where the speaker is indirectly referenced using a noun, for example, *man* in *the man said, "..."*.
- **Deictic anaphora [14] (exophora):** where no explicit indication is made in the text to the identity of the speaker, for example during alternating dialogue between two avatars. In this case no Actor token is identified.

Selecting an avatar as the speaker in the case of direct reference is simply a matter of matching the token with an alias of an avatar in the list. The three cases of anaphora are more difficult, since no alias information is present, and an avatar must be selected based on other information.

3.3.1. Quote Context

A list of avatars present in a scene is required to cater for anaphoric instances since one of these avatars are most likely to be the correct speaker in an anaphoric instance. This list is referred to as the context of the quote, and is implemented using a sorted-list, based on a priority score that may be boosted due to certain conditions, and that degenerates with time. The candidates at the front of the list have the highest scores, and are most likely to be participating in the current scene.

The context is maintained as follows. The fiction text is traversed token by token in order from beginning to end. If a token (or set of tokens) matches an alias of any of the avatar in the avatar list (direct reference), then the avatar is pushed to the top of the list with a score of 1.0. The remaining avatars are demoted by a small amount. Penalties are incurred on the newly promoted avatar's score if the token responsible for the promotion is followed by a possessive ending (for example, *'s* in *Joe* 's) or is marked as a prepositional-complement by the FDG parser (for example *Joe*, in *He said to Joe...*). The operations performed on the scores for each avatar in the context are summarised in Table 1.

Feature	Score Operation
Alias of Character C found	$score_C = 1.0$
Character C demotion	$score_C = score_C * 0.9$
Token indicating C is Possessive Ending	$score_C = score_C - 0.3$
Token indicating C is Prepositional-complement	$score_C = score_C - 0.3$

Table 1: Score operations for avatars in the quote context

3.3.2. Resolving the anaphora

If the Actor token is an instance of anaphora, then a hand-coded decision process is followed to determine which avatar in the quote context should be selected. The decision process is presented graphically in Figure 3, which differs from conventional anaphora resolution algorithms in that it is concerned only with determining the Speaker of an identified Actor, and not all instances of anaphora in the text. Four variables are defined as part of this process as follows:

- ACTOR: refers to the Actor identified in previous processes. This may be blank if the quote has no associated Speech-Verb and Actor tokens.
- BEST: refers to the avatar at the top of the quote context list. This may be blank if the context is empty.
- NEXT: refers to the second avatar in the quote context list. This is used for instances of dialogue between two avatars, where the speaker is implied as a result of the alternation between speaking avatars.
- LAST_SPEAKER: refers to the speaker of the previously handled quote. This may be blank if the current quote is the first to be handled.

The most important factor in this process is the existence of the Actor: either as a direct alias of an avatar, or as an anaphoric reference. In the latter case, gender-matching is used to identify the speaker. Where quotes do not have an associated Actor a two-person dialog is assumed, and an alternation between two avatars is expected. This model could be expanded to handle three way dialogue, but anaphoric instances of this is uncommon in fiction text. The decision process also takes into account errors that may occur in the Actor identification process. For instance, if an Actor is defined, but the corresponding Actor was responsible for the previous quote as well, then BEST is chosen instead.

4. Evaluation of Speaker annotation tasks

The processes for annotating Speech-Verbs, Actors and Speakers in fiction text are evaluated using a corpus of manually annotated fiction books. Each quote in this corpus is annotated with a Speech-Verb and Actor (in cases where the quote is not deictic) and Speaker. Precision and recall values are calculated for the identification of Speech-Verbs and Actors, where precision indicates the percentage of correct annotations with respect to all automatically created annotations, and recall indicates the percentage of correct annotations with respect to all manually created annotations. Accuracy is used to measure the success of the Speaker identification process (since each quote is expected to have a Speaker).

Since Speaker identification is a three-phased process, baseline and cumulative experiments are conducted for Actor

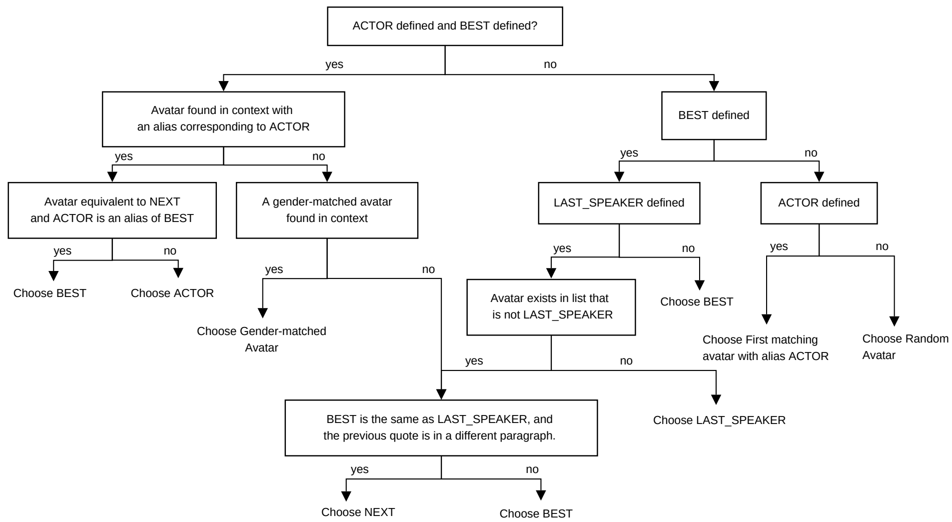

Figure 3: Hand-coded Speaker resolution decision process.

and Speaker annotation. Baseline experiments provide an optimistic indication of precision and recall, by assuming the input data is 100% correct (although experimental error may occur due to incorrect manual annotations). In the case of Actor annotation, the Speech-Verbs from the manual annotation are used as input. Cumulative experiments make use of the automated output of the previous step in order to determine the impact of errors that may have occurred in previous stages. In the case of Actor annotation, the Speech-Verbs annotated automatically are used as input.

The initial salience values used as feature weightings for locating Speech-Verbs and Actors in Section 3 are heuristically defined based on our experience with the manual annotation of the test corpus. Choice of weights represents an experimental factor that can adversely affect any assessment of the Speaker identification process. We measure the effect of this factor by applying a genetic algorithm at each phase to determine optimal salience values that produce the highest recall rates. This provides an assessment of a performance upper bound for the technique which is independent of the parameter values. Since precision and recall are opposing metrics, recall is chosen as a fitness function for the genetic algorithm to ensure that the maximum number of annotations are located correctly. The use of precision to determine fitness would reduce the number of false positives, but also reduce the number of annotations found.

4.1. Results of Speech-Verb annotation

Table 2 presents the overall precision and recall for Speech-Verb annotation over the corpus of 13 fiction books[1]. In total, this corpus contains 9961 quotes that have associated Speech-Verbs, and as the table indicates, 92.90% of the identified Speech-Verbs are correct. However, the salience-based process is less accurate in determining which quotes should not be assigned Speech-Verbs, as is indicated by the low precision rate.

Table 2 also presents the results when using the optimal weightings for each salience feature as determined by the genetic algorithm. Note that both precision and recall are improved by at least 1%. Table 3 presents the salience values for each feature, as determined by the genetic algorithm. The table indicates that for best recall a verb found in an adjacent sentence to a quote must be awarded the highest salience. The table also indicates standard deviation over the 13 books regarding these values, indicating that there is overlap in some of the optimal salience weights across different books, but not between the lowest scoring and highest scoring features.

	Man.	Aut.	Corr.	Prec.	Recall
Baseline	9961	11090	9254	83.44%	92.90%
Weighted	9961	11090	9404	84.80%	94.41%

Table 2: Results of Speech-Verb annotation using baseline and weighted salience schemes over corpus of 13 books.

Feature	Mean Salience	Std. Dev.
Main Verb	+3.28	1.80
Hypernym	+4.86	2.32
Verb in adjacent sentence	+7.69	2.01
Proximity (prev. sent.)	+2.43	1.81
Proximity (next sent.)	+2.85	2.03

Table 3: Optimal salience values for Speech-Verb annotation, determined by genetic algorithm.

Figure 4 illustrates the precision and recall values of the weighted Speech-Verb annotation process across the 13 fiction books of the corpus. The books are arranged in order of ascending sentence length (and readability) from left to right. The figure illustrates that there is little variation in recall across the different books, while precision fluctuates dramatically. There is also no evidence to suggest that the task becomes more difficult as the sentence-length increases.

The above results indicate that the Speech-Verb of a quote may be found with very high recall using a salience based method. The identified features play an important role in the

[1]Columns in all tables correspond to the following key: **Man.** indicates the number of manual annotations; **Aut**. indicates the number of annotations made automatically; **Corr.** indicates the number of correct automatic annotations; **Prec.** indicates the precision.

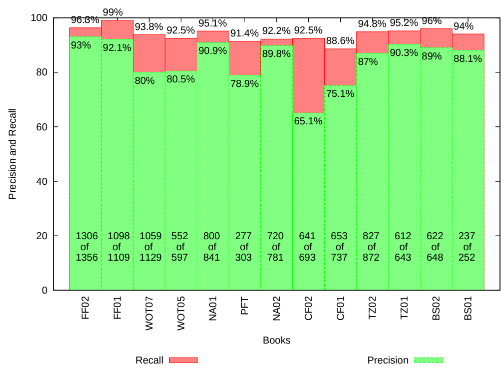

Figure 4: Precision and recall of Speech-Verb annotation across 13 books

location of the verb, and the appropriate choice of weightings for these features are necessary for optimal recall values.

4.2. Results of Actor annotation

Precision and recall rates observed for the process of locating the Actor of a quote are presented in Table 5. The baseline process results in a recall rate of 86.59%. As expected this result is higher than the cumulative recall of 85.10%, since the input Speech-Verbs are correct. This indicates that the correct annotation of the Speech-Verb has an impact on the Actor identification process. However, the use of optimal weightings for the Actor features results in recall which is higher even than the baseline, indicating that error in Speech-Verb annotation can be offset through the use of optimised weightings.

Table 4 presents the optimal salience values determined for each feature used to locate the Actor of a quote. Surprisingly, the salience values produced are remarkably similar, all in a range of between 4.00 and 5.69, and all with a standard deviation of greater than 2.20. However, as Table 5 indicates, these weightings produce an increase in recall of more than 3%. Precision, however is not increased using these weights, which is expected since the genetic algorithm employed makes use of only the recall value in its fitness function.

Feature	Mean Salience	Std Dev.
Abbreviation	+4.51	2.76
Noun	+4.41	2.21
Proper Noun	+4.00	2.49
Pronoun	+5.08	2.80
Subject	+5.24	2.86
Object	+5.69	2.50
Distance from verb (penalise)	-0.51%	0.26

Table 4: Optimal salience values for Actor annotation, determined by genetic algorithm.

Individual precision and recall rates for each book in the corpus are illustrated in Figure 5, where books are arranged in order from left to right based on the sentence length (readability). The worst performing books in the corpus in terms of recall are the NA02 and the CF01 books. Poor results for the NA02 book is explained by the use of multiple speech quotes in a single sentence, where the location of Actors becomes becomes more ambiguous since the number candidates per quote

	Man.	Aut.	Corr.	Prec.	Recall
Baseline	9955	9548	8620	90.28%	86.59%
Cumulative	9955	10488	8472	80.78%	85.10%
Weighted Cumulative	9955	10879	8782	80.72%	88.22%

Table 5: Results of Actor annotation using baseline and weighted salience schemes over corpus of 13 books.

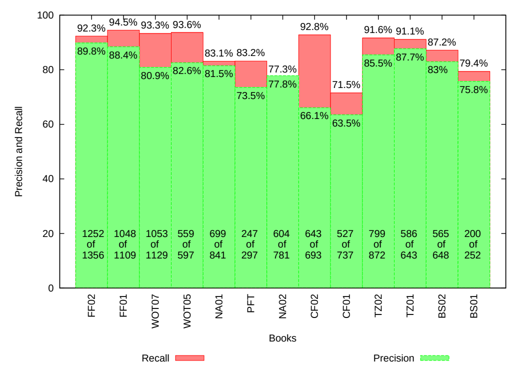

Figure 5: Precision and recall of Actor annotation across 13 books

increases. Poor results in the CF01 book are explained by the use of titles, for example, *Brother* in “Brother Jerome” and “Brother John”. The actor annotation process incorrectly selects “Brother” as the Actor instead of the avatar name that follows it.

The Actor annotation process is shown to produce high recall rates given appropriate weightings for the different salience values. Precision values are markedly lower than the Speech-Verb task however, and stand to be improved with tuned weightings, a necessary task since incorrectly assigning Actors to quotes that should not have Actors negatively affects the Speaker identification process.

4.3. Results of Speaker annotation

Accuracy is used to evaluate Speaker identification, since every speech quote is expected to have a speaker, and therefore precision and recall become equivalent. Accuracy for the Speaker identification process is presented in Table 6, in which the baseline experiment scores an accuracy of 81.71%. As with Actor annotation, the cumulative experiment results in reduced accuracy of 79.4% since accumulated errors from Speech-Verb and Actor identification adversely affect the Speaker identification process. Note that each Avatar has approximately two manually defined aliases.

The majority of the error in the Speaker annotation process is attributed to the context model which relies on the quality of aliases defined in the manually created avatar list. The context model, while accurate in many instances, does not reliably select the correct avatar. This is a result of the quality of the aliases defined in the avatar list: if an alias encountered in the text is not defined for a specific avatar, then that avatar will not appear correctly in the context model. Conversely, the over-definition of aliases may result in more than one avatar having similar aliases, confusing avatar selection. Given these problems the accuracy in Table 6 stands to improve given fur-

	Manual	Correct	Accuracy
Baseline	12221	9986	81.71%
Cumulative	12221	9704	79.40%

Table 6: Results of Speaker annotation using baseline and cumulative schemes over corpus of 13 books.

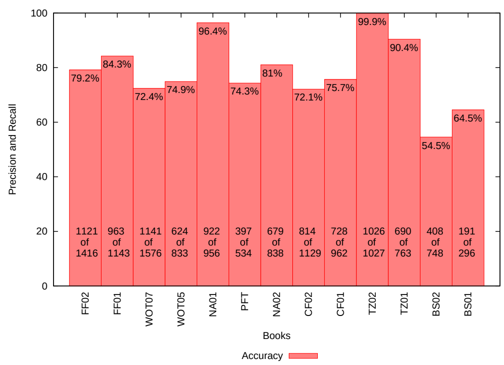

Figure 6: Accuracy of Speaker annotation over corpus of 13 books

ther research into the construction of more accurate context and aliasing techniques.

Figure 6 illustrates the accuracy rates for each individual book in the corpus. The worst accuracy is achieved over the BS02 book, explained by a narrative which only introduces a primary avatar very late in the book, a technique for which the context model is not designed. However, the Speaker annotation process is shown to be extremely successful, especially in the instance of the TZ02 book, obtaining over 99.9% accuracy.

The results presented in this section indicate that Speaker annotation may be achieved at an average rate of 81.71% accuracy, dependent on the type of narrative used in the book. It is expected that improved results may be achieved through the development of a more complex context model, as well as avatar aliasing scheme.

From Table 6 it is expected that on average, only 20.6% of a fiction book annotated using this method would require manual correction. This represents a substantial saving in effort over the alternative of hand annotating the entire text from scratch, especially since the process used is performing the task in the absence of any genre specific training data.

5. Conclusion

Speaker identification is shown to be feasible through the use only of a salience-based annotation process that does not make use of complex machine learning or logic-based techniques. Contributions include the use of a salience-based technique for script extraction, a process that has not been previously investigated. The primary contribution is a method for creating dramatic scripts representing dialogue in fiction text which can serve as the basis for automatic generation of virtual environments.

Future work includes the development of other tools to extract additional information from text to be used to generate visual representations of the text.

6. References

[1] Nina Wacholder, Yael Ravin, and Misook Choi, "Disambiguation of proper names in text," in *Proceedings of the fifth conference on Applied natural language processing*, San Francisco, CA, USA, 1997, pp. 202–208, Morgan Kaufmann Publishers Inc.

[2] William W. Cohen and Sunita Sarawagi, "Exploiting dictionaries in named entity extraction: combining semi-markov extraction processes and data integration methods," in *KDD '04: Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, New York, NY, USA, 2004, pp. 89–98, ACM Press.

[3] Kevin Glass and Shaun Bangay, "Hierarchical rule generalisation for speaker identification in fiction books," in *Proceedings of SAICSIT '06*, Republic of South Africa, 2006, pp. 31–40.

[4] Jason Zhang, Alan Black, and Richard Sproat, "Identifying speakers in children's stories for speech synthesis," in *Proceedings of EUROSPEECH 2003*, Geneva, September 2003.

[5] Christopher Kennedy and Branimir Boguraev, "Anaphora for everyone: Pronominal anaphora resolution without a parser.," in *COLING*, 1996, pp. 113–118.

[6] Ruslan Mitkov, Richard Evans, and Constantin Oruasan, "A new, fully automatic version of mitkov's knowledge-poor pronoun resolution method," in *Proceedings of the Third International Conference on Intelligent Text Processing and Computational Linguistics (CICLing-2002)*, Mexico City, Mexico, February, 2002.

[7] Marin Dimitrov, "A light-weight approach to coreference resolution for named entites in text," M.S. thesis, University of Sofia, February 2002.

[8] Breck Baldwin, "CogNIAC: high precision coreference with limited knowledge and linguistic resources," in *Proceedings of the ACL'97/EACL'97 workshop on Operational factors in practical, robust anaphora resolution*, Madrid, Spain, 1997, pp. 38–45.

[9] Kevin Glass and Shaun Bangay, "Evaluating parts-of-speech taggers for use in a text-to-scene conversion system," in *Proceedings of SAICSIT '05*, Judith Bishop and Derrick Kourie, Eds., White River, South Africa, September 2005, pp. 20–28.

[10] Connexor, "Connexor machinese syntax," 2007, http://www.conexor.fi/ [Accessed on 25 September 2007].

[11] Paola Velardi, Paolo Fabriani, and Michele Missikoff, "Using text processing techniques to automatically enrich a domain ontology," in *FOIS '01: Proceedings of the international conference on Formal Ontology in Information Systems*, New York, NY, USA, 2001, pp. 270–284, ACM Press.

[12] Christiane Fellbaum, Ed., *WordNet: An Electronic Lexical Database*, MIT Press, 1998.

[13] Kevin Atkinson, "12dicts," 2003, http://wordlist.sourceforge.net/ [Accessed on 22 November 2005].

[14] Yusuke Shinyama, Takenobu Tokunaga, and Hozumi Tanaka, "¨Kairai¨ - software robots understanding natural language," in *Third International Workshop on Human-Computer Conversation*, 2000.