Sequence Package Analysis: A New Natural Language Understanding Method for Performing Data Mining of Help-Line Calls and Doctor-Patient Interviews

Published Proceedings of the Natural Language Understanding and Cognitive Science Workshop at the 6th ICEIS (University of Portugal, April 13, 2004)

Amy Neustein, Ph.D.

Linguistic Technology Systems
1055 River Road, Suite 1013
Edgewater, New Jersey 07020, U.S.A.
lingtec@banet.net

Abstract. Designers of audio mining programs must confront the complexities of natural language dialog, which is replete with ambiguities, circumlocutions and ellipses. Speakers often make requests, lodge complaints, or report on problems in such roundabout ways that attempts to find a statistically probable word match between the application vocabulary and the user's speech can yield unsatisfactory results. One solution to this problem is the application of Sequence Package Analysis (SPA) to perform intelligent data mining. The purpose of this paper is to demonstrate how Sequence Package Analysis, as a new AI-based natural language understanding method, can be used to gather important business intelligence by learning what customers are really complaining about when they make contact with customer call centers. The paper also looks at how SPA can be used in mining tape recordings of doctor-patient interviews to uncover important medical history data, often buried in the ambiguity of patient dialog. The paper, however, not only looks at the applications of SPA (e.g., industrial and medical), but demonstrates that SPA, as a new data mining technique, is an innovative heuristic approach to data mining that can bring machines one step closer to understanding humans.

1. Introduction

Audio data mining is an exciting and rapidly developing field that draws upon a variety of methodological approaches. These approaches range from pure natural processing (NLP), requiring the manual development of complex rules and algorithms, to somewhat less knowledge-intensive techniques derived from machine learning and information retrieval (IR) that may use, for example, an advanced statistically-based IR method – Latent Semantic Indexing (LSI) – to create a term-document matrix to which Singular Value Decomposition (SVD) is applied to find semantic relations among sets of words (Chu-Carroll and Carpenter, 1999; Reichl et al., 1998).

Recently, a new AI-based natural language understanding method, known as Sequence Package Analysis (SPA), has emerged in some thought-provoking discussions of data mining approaches that can be used to analyze the quality of service in call centers. Paprzycki et al. (2004), in reviewing the literature dealing with varied methodological approaches to mining call center data, point to Sequence Package Analysis as a method that can "caption the text to which data mining is applied" so as to enable the "capturing of early warning signs of caller frustration."

In this paper, the author demonstrates how SPA can be used to mine large amounts of audio data to learn sensitive information about "events." ("Event" is a commonly used term in information extraction/retrieval to refer to things such as who did what, when and where, or to refer to the kinds of business/organizational relationships among the principals involved.) In particular, the kinds of sensitive information that can be learned by the SPA method of intelligent data mining are the *subtleties* surrounding those events. The author shows that these "subtleties" are not mere barometers of customer satisfaction; instead, they furnish crucial business intelligence – what are customers really complaining about, and are they being properly serviced? Such business intelligence data is indispensable to an enterprise wanting to deliver the best possible products and services.

In addition to call centers, the author looks at doctor-patient interviews as another area where effective data mining is needed. In the medical interview, as in the call center, the goal of SPA is to do more than just measure satisfaction variables. That is, tape-recorded interviews of doctor-patient interactions are mined for the purpose of finding important medical history data, often buried in the ambiguity of patient dialog. This yields important information that is not revealed by simply looking for patient satisfaction criteria (Neustein, 1989, 1999, 2001).

2. Methodology and Related Research

Sequence Package Analysis, like other methods of information extraction and retrieval, enables automated knowledge capture by finding interesting patterns from data sets. SPA shares the practical edge such methods enjoy over manual methods that require tremendous human capital. But where SPA departs from other data mining methods, such as Latent Semantic Indexing, is that SPA's primary analytical focus is the *unit of interaction* in its entirety, referred to as a "sequence package," rather than a single or multiple lexical item, such as a "content word" (e.g., "servicing") and its corresponding "content term root" (e.g., "service").

SPA focuses on the sequence package – defined as a series of related turns and turn construction units (lexically bounded parts of turns at which point the turn may conceivably, but not necessarily, be yielded to the next speaker) that are discretely packaged as a sequence of conversational interaction. This is done in order to accommodate better to how people really talk. In the current speech systems, when speakers do not produce the syntactically expected "keyword," or its closely related morphological variant, the system cannot perform a useful word match. This is a serious shortcoming, particularly when one adds in complicating emotional and behavioral factors. In such cases, the chance of a speaker uttering the expected key word diminishes sharply. This can easily occur when a caller to a help-line desk is in a high state of distress; or when a patient is in a state of turmoil and confusion after having just learned of an unfavorable diagnosis; or when a terrorist is frantically trying to arrange an attack while at the same time trying very hard to conceal his plans and operations.

SPA adjusts to these many factors that make speech less than perfect, offering grammatical constructs to data mining programs that better conform to the patterned way speakers really talk, with all the ambiguities, ellipses, and other imperfections of natural language dialog. What SPA locates in a corpus of data are the discrete patterns underlying what may appear to most systems as unwieldy talk. Unfortunately, much of this data is discarded or ignored by conventional speech systems, as such kinds of data appear far too amorphous to be useful to the speech recognizer.

To map out the discrete patterns of natural language dialog so that mining programs can adjust to the rich cadences of natural speech – replete with vagueness and imprecision, in addition to metaphors, idioms, shibboleths, and other forms of colloquialisms – SPA draws mainly from the field of conversation analysis, a discipline whose origins are in sociology.

Conversation analysis breaks down natural language communication into its primary units of analysis: sequences and turns within sequences (rather than isolated sentences or utterances). This allows for the study of how participants in a dialog organize their verbal interactions in both mundane conversations and in institutional talk, such as doctor-patient interaction. In this way, conversation analysts have been studying interactive dialog for over 35 years as a socially organized activity. In essence, the conversation analyst can be distinguished from the linguist in that the linguist focuses on grammatical discourse structure, while the conversation analyst focuses on social action (McIlvenny and Raudaskoski, 1992).

The empirical methods used by conversation analysts include the recording and transcription of verbal interactions by using highly refined transcription signals to identify both verbal components and paralinguistic features, such as stress, pauses, gaps, overlaps and changes in intra-utterance spacing (Atkinson and Heritage, 1984, pp. ix-xvi). The most valuable research findings from the field of conversation analysis have launched the development of SPA as a new AI-based natural language understanding method.

There are two ways that an SPA-driven mining program can work. First, it can serve as an “add on” layer for conventional data mining programs that are built on vector-based models. Then the “global weighting” to be applied for the next layer of analysis need not be limited to content words or their term roots; rather, it could also encompass sequence package material.

Second, SPA might be used as a wholly integrated system rather than as an “add on” layer to conventional IR systems. In such a case, data mining programs would use sequence package grammars rather than content words as their starting point. Such a use would allow the building of a vocabulary of appropriate content words, and their corresponding root terms, without necessarily having to have an *a priori* knowledge of such words. Using this same heuristic approach, a data mining program would seek to discover, in addition to content words and their term roots, new or related sets of sequence packages that demonstrate the patterned way humans engage in interactive dialog. In the last analysis, such discoveries made about how natural language dialog really works can only serve to bring machine learning one step closer to truly understanding humans.

3. Call Centers

One of the greatest assets to an enterprise is the amassing of good business intelligence. Today, call centers provide a fertile environment for culling such intelligence: with an emphasis on quality assurance, enterprises record thousands of hours of calls between customer service agents and callers, thus creating a gold mine of data for gathering business intelligence about the customer or client. Yet, little of this fertile resource of customer data is ever sufficiently mined to learn useful information about customers' needs and preferences, which can be very important in making decisions about which customers might be receptive to an expanded line of products and services.

But even more important, enterprises are not getting the feedback that is essential to preserve an existing customer base by learning from customers' calls into help-lines what is the gravamen of their complaints. And although customer relationship management is crucial to the viability of an enterprise, the harsh reality is that when customers fail to use the expected keywords in articulating their complaints and requesting assistance, this poses a major problem for standard data mining programs, making it very hard for such mining programs to uncover critical information about the nature and frequency of customer complaints.

Data mining programs that use SPA could surmount this common difficulty. Here are a couple of reasons why SPA is especially well suited to the difficult task of uncovering customer complaints when callers use words outside of the application vocabulary. First, the subtleties of dialog, as evident in the indirect ways caller often express themselves, can be detected by SPA, which can pierce through layers of obfuscated dialog to uncover the essence of customers' complaints to help-line desks. This is something conventional systems cannot do easily, because they can only spot words and phrases (or their morphological equivalents) to see if they conform to the speech system's application vocabulary. Second, because help-line dialog is highly scripted (where human agents more or less read from a well formatted script), an SPA-driven system can find the relevant sequence package by looking for its proximity to the agent's own scripted text (Neustein, 2002). Again, this is something that conventional speech systems would not be able to do readily, because the lexical items found in close proximity to the agent's scripted text might have such a poor statistical word match to the speech system's application vocabulary that not much sense could be made of the dialog anyway.

3.1. Demonstration

Here is an example of a help-line desk call from a caller who is distressed over his inability to get his scanner to work. Because the caller is distressed he lapses into a repetitive use of vague descriptors to explain his problem and equally vague and elliptical descriptors to request assistance. It can easily appear that he seems to be "lost" in his own world in which no one else can actually understand what he is trying to say. However, by applying SPA, a data mining program can begin to make sense out of what this distressed caller is trying to say because his talk, albeit circuitous and elliptical, is understood by distinct conversational sequence patterns that can be detected by SPA grammars.

Agent: "How can we help you?"

Caller: "I really can't do this myself. I can't get this to work without someone coming here. I really don't know what to do with this."

The conversational sequence pattern that can be seen here is the caller's repeated use of pronouns (and similar unnamed referents) – standing in place of noun referents to identify the problem and make a proper request for assistance – in close proximity. Notice that the caller never mentions "service call" or "technician" in his request for help. To learn what the caller is referring to when the caller fails to utter the expected keywords, an SPA-driven data mining program would get to the root of the customer's complaint. Here's how: First, it would map out the sequential patterns of pronoun usage, identifying in this sequence package a high use of pronouns in close proximity. The sequence package is divided into three parts:

1) a short, condensed complaint – referenced by pronouns;
2) an expansion of the complaint – embedded within a series of pronouns in place of nouns; and
3) a recycling of a short complaint – referenced by pronouns.

The pronouns and (unnamed referents) appear here in bold italics:

I really can't do ***this myself.***

I can't get ***this*** to work without ***someone*** coming ***here***.

I really don't know what to do with ***this.***

After finding the complaint template, the mining program would use a heuristic procedure for uncovering what the caller is actually requesting, by piercing through its vagueness and non-specificity. To do this, the program would look at the concatenation of lexical items and flag the *second* of the three parts of this complaint sequence template, which is where the "expansion" of the complaint can be found. What would be uncovered is the caller's need to have "someone coming here." After that, the semantic meaning of this statement ("someone coming here") would be made transparent to the intelligent mining program working with speech recognition systems whose probabilistic knowledge of conceptual relationships would dictate that in a call to a help-line desk the caller who refers to "someone coming here" is not offering an invitation to dinner, but rather asking for a service call.

Finally, after the identification of the nature of the caller's complaint and his request for assistance, the intelligent data mining program would classify this call accordingly: a complaint about a malfunctioning scanner and a simultaneous request for a service call. In so doing, the program would add the phrase "someone coming here" as a non-keyword correlate to "service call" or "technician." Consequently, each time the program identified a non-keyword correlate to the expected keyword, the lists of non-keywords, representing the panoply of subtle and roundabout ways customers express themselves, would expand. The end result would be the creation of an application vocabulary that better reflects the dialog patterns of customers so that call center recordings can be mined more effectively for the purpose of gathering useful business intelligence about customers' complaints.

4. Doctor Patient Interviews

Medical history-taking provides the physician with one of the richest resources for making accurate diagnoses and instituting proper treatment. And in our cost-conscious world, medical histories take on an even greater value: by getting a complete picture right at the beginning, doctors can cut back on the many unnecessary and costly tests that are given to patients to provide the clinician with the necessary information that the patient failed to provide in his history. But ironically, what should be a relatively simple task is often complicated by so many conversational difficulties that the value of the history is minimal. Several things contribute to these conversational problems.

First, doctors have large loads and very little time to spend with each patient. Working under such constraints, it becomes hard to listen to each patient, especially to those who are long winded. Second, patients all too often present their medical problems and histories in such a roundabout way that they leave their doctors with more questions than answers. Third, patients increasingly change providers, as their health plans are determined by their jobs, marriage/divorce, relocation and other demographic factors. As a result, their care providers don't have the opportunity to really get to know them as the old fashioned doctors once knew their patients.

One solution to these difficulties is intelligent mining of tape recordings of doctor-patient interviews. These are used to uncover important medical history data that escapes the practitioner's attention, especially when patients are circuitous and vague. Intelligent mining can also uncover the emotive content of the medical interview in which patients sometimes express their fears and concerns in such indirect ways that it escapes the notice of the physician altogether. Such mining can be done in real time during the medical visit or afterwards, so that feedback can be given to the doctor for a follow-up visit. Either way, intelligent data mining can extract important diagnostic data that is submerged beneath layers of dialog. In addition, it can uncover the patient's fears and worries. The payback would then be twofold: 1) more efficient handling of the case; and 2) an improvement in patient satisfaction.

4.1 Illustrations

Here are two examples of doctor-patient dialog, derived from tape-recorded interviews of patients at a New York teaching hospital. In both examples, the patients give very important history data to the physician at the wrong place in the medical encounter, when the doctor's attention has been diverted to something else. The patients, despite ample opportunity to give this vital history data when the doctor was looking for it, withheld mention of these important facts, for whatever reason, until later in the medical encounter. In the first example, it is family medical history that is initially withheld; in the second case, it is the patient's chief complaint (the primary reason the patient is seeking medical care) that is withheld until later on in the interview.

Whether or not we can determine why patients defer such critical history data (fear of a serious diagnosis may play a role), it is beyond question that such delayed disclosures – no matter how critical to formulating an accurate diagnosis and an effective treatment plan – can go over the doctor's head. Inasmuch as it is the doctor and not the patient who directs the questioning in the medical interview (Frankel, 1990), it is the doctor who likewise directs the "listening" in the medical interview. As a result, doctors can selectively "tune out" at times, particularly when patients give information at the wrong place in the interview (Neustein, 1989, 2001).

Case One

Patient: "I become terribly worried about my pain, which reminds me of the arthritic pain that my sister had, which turned out to be bone cancer, so I worry whenever I have pain because I don't know if it is what she had."

In this case, the patient made this revelation about a family history of bone cancer when the physician was busy conducting his physical examination. Ironically, a few minutes earlier, when the doctor during the medical interview had asked specifically about any history of family illnesses, the patient was silent. However, one must look closely at how the patient actually relayed this information about her family history. What is striking is that when the patient revealed the family history, she obfuscated her overriding medical concern – which was whether her own pain could reflect an undiagnosed case of osteosarcoma (bone cancer) – by embedding it in a series of narrative type statements, as if she were telling a story to a neighbor or friend. By revealing this medical history information in the form of a *narrative*, and at a juncture in the examination where medical history questions were no longer topical, the patient possibly hoped her fear of bone cancer would be given less attention than if she came out with it directly when asked about her family history. Yet, at the same time, the fact that the patient brought up her sister's bone cancer at all indicates that she wanted the doctor to address her concerns and to possibly allay her fears.

Although this patient straddled the fence, using dialog that conveyed the importance of this history datum while at the same time trying to downgrade its impact, SPA could have retrieved from the recording of this interview this otherwise "lost" data (and its strong emotional implications for the patient).

Here's how. First, SPA would map out the sequential patterns of narrative usage, identifying in this sequence package a high usage of narrative phrases in close proximity. The sequence package is divided into four parts:

1) a short, condensed and somewhat nonspecific concern preceded by a narrative phrase;
2) an expansion of the concern, citing the troublesome datum ("bone cancer"), which is embedded within two narrative predicates;
3) a recycling of the nonspecific concern preceded by a narrative phrase; and
4) a reference back to the expanded concern, but only with the use of pronouns that serve as anaphors, referring back to the expanded concern.

The narrative predicates appear here in bold italics:

I become terribly worried about my pain
which reminds me of the arthritic pain that my sister had ***which turned out to be*** bone cancer
so I worry whenever I have any pain
because I don't know if it is what she had

Second, upon finding sequences that are punctuated by a high amount of narrative predicates, pronouns, and anaphors, the SPA-driven mining program would require a heuristic procedure for discerning from this vague and winding dialog the true nature of the patient's complaint. To do that, SPA grammars would search for the specific content material embedded within the two narrative predicates appearing in the second part of the four-part sequence package **("which reminds me of…which turned out to be…")**. In addition, such grammars would uncover specific content words ("worried") and their term roots ("worry") upon locating this narrative sequence package in the patient's dialog. At the end, the SPA mining program would offer the physician the information that this patient has bone cancer in her family and that she is quite concerned that she too might be stricken with bone cancer, just as her sister was. Such information might help the physician to make the decision to order periodic bone scans and MRIs so as to keep a careful watch for the possible development of osteosarcoma, a disease that may be more successfully treated if caught in its early stage(s) of development.

Case Two

Doctor: "Was it a dull achy pain?"
Patient: "Perhaps, maybe."
Doctor: "But it wasn't a sharp pain or a squeezing pain, was it?"
Patient: "No."
Doctor: "Just like a dull achy pain. And as you stated previously, you said that when you lay down it came again and when you got up it went away?"
Patient: "Yes."
…
(Later in the interview)
Doctor: "Do you ever have heartburn or indigestion?"
Patient: "I had trouble with my stomach. I ate something today and I think it didn't agree with me, so I had pain? [Question mark indicates patient is questioning herself] Uh I mean just heaviness…so I came here."

In this case, the patient began the interview by characterizing her complaint as a "dull, achy pain" around her heart. Much later in the interview, however, while the doctor asked rather perfunctory questions about neurological, digestive, musculo-skeletal and other organ systems, the patient revealed that the problem necessitating her trip to the hospital was not heart-related at all, but rather digestive in origin, referring to her problem as "just heaviness." But because this chief complaint of gastric disturbance was made so late in the interview, and at a juncture when the chief complaint was no longer the focus, the doctor gave it little or no attention.

Proceeding from the assumption that the patient's primary complaint was heart-related, not digestive, the doctor ordered a full set of cardiac tests: EKG, stress test, and blood analysis. Had the physician been attuned to the patient's chief complaint of "heaviness," he would likely have pursued a different diagnostic course altogether, testing her for a hiatal hernia or GERD (gastroesophageal reflux disease) or some other form of gastric pathology. Sadly, this patient's medical concerns were not adequately handled.

A closer look at the doctor-patient dialog reveals that the patient appeared ambivalent about her symptoms until the interview was nearly complete, when she showed a bit of confidence and certainty in revising her chief complaint, claiming it was a gastric problem and not a heart problem. But even in her efforts to appear certain, she preserved the ambiguities that punctuated her earlier dialog. This can be seen in her upward querying inflection in which she questioned the validity of her own symptom descriptor – **"so I had pain?"** – before offering a revised version of her symptoms – **"Uh I mean just heaviness."**

This doctor was confronted with a patient who continually cloaked her symptom descriptions in ambivalent terms. He had no choice but to fill in the gaps for her. And since she did arrive at the hospital emergency room, rather than a doctor's office, the physician could not afford to overlook the more serious diagnosis, which would have been cardiac in nature. To meet these challenges, the doctor quickly got a handle on the interview; and in so doing, converted her ambivalent descriptors into ones with greater certainty. As a result, he made a diagnosis of cardiovascular disease. Unfortunately, by the time she became somewhat more assertive – stating later on in the interview that she believed her problem was gastrointestinal (presumably caused by something she ate) – it was already too late to be taken seriously by the doctor who had struggled through an interview with an ambivalent patient.

Applying a discrete set of SPA grammars to this dialog, designed to locate the conversational sequence patterns of alternating equivocal and certainty descriptors, would identify and bring to the physician's attention the lost or buried history data.

Here's how. First, an SPA program would map out the sequential patterns of "equivocality" versus "certainty," identifying in this sequence package the frequent alternation between ambivalence and certainty. The sequence package is divided into four parts:

1) Patient – an equivocal (confirmation of a) descriptor;
2) Doctor – an unequivocal contrastive descriptor to the prior equivocal one;
3) Doctor – a compression of an equivocal descriptor with a certainty descriptor;
4) Patient – a recycling of the equivocal descriptor, succeeded by a certainty descriptor.

The equivocality versus certainty descriptors appear in bold italics:

Doctor: Was it a dull achy pain?

Patient: ***Perhaps, maybe***.

Doctor: ***But it wasn't a sharp pain or a squeezing pain, was it?***

Patient: No.

Doctor: ***Just like a dull achy pain. And as you stated previously, you said that when you lay down it came back again and when you got up it went away?***

Patient: Yes.

Doctor: Do you ever have heartburn or indigestion?

Patient: ***I had trouble with my stomach. I ate something today and I think it didn't agree with me, so I had pain? [Question mark indicates patient is questioning herself] Uh I mean just heaviness…so I came here.***

Second, upon finding sequences that are punctuated by a repetition of equivocality versus certainty descriptors, the SPA-driven mining program would require a heuristic procedure for discerning from this constant alternation of ambivalence and certainty what is the true nature of the patient's chief complaint. But unlike the prior examples discussed above, the sequence packages here cannot be confined to one or two speaking turns. Instead, sequence package grammars must be able to expand flexibly so as to accommodate large episodes of talk where equivocality versus certainty is found. In this example, it was at the end of the medical interview where the patient sought to amend her chief complaint. A grammar would therefore have to be expansive enough to stretch from the early part of the interview, where the dual features of equivocality versus certainty were located, to the end of the interview, in which this same contrastive pair type emerged once again as the patient revised her chief complaint. In the last analysis, the wide range of subtleties endemic to dialog can be better understood by using grammars that can search through large episodes of talk. In this case, SPA could make the patient's habitual ambivalence, and her failed attempt to revise an important piece of medical history, more visible to a data mining program. And in all such cases, unraveling the subtleties of doctor-patient dialog affords doctors one of the richest sources of diagnostic data.

5. Conclusion

As audio data mining becomes increasingly important to the legal and medical professions, the broadcast news industry, and the industrial sector, better methods are needed to prefect its performance capabilities. Sequence Package Analysis, as a new AI-based natural language understanding method, offers a set of flexible grammars that can parse small segments or an entire dialog to detect patterns that are indigenous to the dialog. By learning such patterns, depicted in the form of sequence packages, the system can automatically mine crucial business intelligence data that would have otherwise required costly and time-consuming human analysis.

SPA brings a new method of parsing dialog to the field of data mining, with the result that important information submerged in the flow of natural speech, with all its imperfections and ambiguities, can be properly mined. No matter how SPA is implemented in mining programs, its heuristic approach to understanding spontaneous speech as a socially organized activity offers a new way to retrieve information that may otherwise be compromised by the natural obfuscations of speech.

References

Atkinson, J.M. and Heritage, J. (1984). Transcript notation. In J.M. Atkinson and J. Heritage (Eds.), *Structures of Social Action: Studies in Conversation Analysis*. Cambridge: Cambridge University Press, pp. ix-xvi.

Chu-Carroll, J. and Carpenter, B. (1999). Vector-based natural language call routing. *Computational Linguistics*, *25*(3): 361-388.

Frankel, R. (1990). Talking in interviews: A dispreference for patient-initiated questions in physician-patient encounters. In G. Psathas (Ed.), *Interaction Competence*. Washington, D.C.: University Press of America, pp. 231-262.

McIlvenny, P. and Raudaskoski, P. (1992). The mutual relevance of conversation analysis and linguistics: A discussion in reference to interactive discourse. In L. Heltoft and H. Haberland (Eds.), *Proceedings of the Thirteenth Scandinavian Conference on Linguistics*. Roskilde, Denmark: Department of Languages and Culture, Roskilde University, pp. 263-277.

Neustein, A. (1989). Medical history-taking as an interactive event. In W. von Raffler-Engel (Ed.), *Doctor-Patient Interaction*. Amsterdam/Philadelphia: John Benjamins Publishing Co, pp. 61-76. Pragmatics and Beyond New Series, Vol. IV.

Neustein, A. (1999). How “sequence packages” can aid language understanding. *Speech Technology*, *4*(4): 36-37.

Neustein, A. (2001). Using sequence package analysis to improve natural language understanding. *International Journal of Speech Technology*, *4*(1): 31-44.

Neustein, A. (2002). “Smart” call centers: Building natural language intelligence into voice-based applications. *Speech Technology*, *7*(4): 38-40.

Paprzycki, M., Abraham, A., and Guo, R. (2004). Data mining approach for analyzing call center performance. *The 17th International Conference on Industrial and Engineering Applications of Artificial Intelligence and Expert Systems*. Ottawa, Canada. In *Lecture Notes in Computer Science*. Germany: Springer Verlag.

Reichl, W., Carpenter, B., Chu-Carroll, J. and Chou, W. (1998). Language modeling for content extraction in human-computer dialogues. *The 5th International Conference on Spoken Language Processing. Sydney, Australia*