

# Sequence Package Analysis: A New Method for Intelligent Mining of Patient Dialog, Blogs and Help-line Calls

Amy Neustein, Ph.D.

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

[lingtec@banet.net](mailto:lingtec@banet.net)

**Abstract**—The ambiguities, repetitions and ellipses commonly found in natural language dialog continue to hinder speech (and text) analytic mining programs that glean business intelligence data from consumer help-line calls, or extract important medical diagnostic information from doctor-patient interviews or consumer-generated health-related blogs. This poses an even greater problem when such mining programs attempt to extract critical emotional data from natural language dialog. At present, speech (and text) analytic programs that mine natural language dialog for signs of distress, frustration, anger or other human emotions are still largely ineffective, because conventional speech systems that are limited to a set of key words and phrases cannot process speech as it actually occurs; if a speaker or blogger fails to use the word(s) found in the speech application’s vocabulary, the program yields a poor statistical word match (or no match). This paper shows how Sequence Package Analysis is informed by a set of algorithms – representing some of the more complex semantic aspects of communication in addition to syntax – that can interpret less than perfect natural speech, enhancing intelligent mining of recordings of doctor-patient interviews, customer care help-line calls, and consumer-generated health-related blogs.

**Index Terms**—blogosphere, call centers, customer-care help-line calls, data mining, doctor-patient dialog, medical interviews, natural language understanding, sequence package analysis, speech analytics

## I. INTRODUCTION

*Sequence Package Analysis* offers a new and advanced Natural Language Understanding (NLU) method whose parsing structures (consisting of context-free grammatical units with notations for related prosodic features, such as elevated inflection or pitch or changes in rate of speech) are designed to reflect, in addition to syntax, some of the more complex and ambiguous *semantic* aspects of communication. A discussion of the methodological basis of Sequence Package Analysis (SPA), formulated by this author, first appeared in the *International Journal of Speech Technology* [1]. SPA has been referred to by AI researchers as among the more advanced parsing methods for “captioning the text to which data mining is applied” so as to better detect the subtleties of human emotions [2]. By relying on the *entire* sequence package – a series

of related speaking turns and parts of turns – rather than on individual words or combination of words and their attendant stress patterns, SPA can enable a speech system to detect, among other things, the wide range of speakers’ emotions found in doctor-patient recorded dialog, customer-care help-line calls, and consumer-related blogs.

Because blogs constitute a rapidly expanding medium of communication, with tens of thousands of blogs created each day [3], blog experts, such as Technorati’s founder and CEO, David Sifry, estimate that the blogosphere will double in size once every six months [4]. While there is still no consensus among experts over the interconnectivity of blogs (some have found “a majority of blogs link sparsely or not at all with other blogs” or connectivity itself to be “overrepresented” by the “privileged subset of popular blogs” known as the “A-list”) [5], there is no dispute that blogs closely resemble the informality of conversational dialog, as opposed to the structured format of written text, and can thereby benefit from advanced NLU methods.

Experts in the field of text analytics have acknowledged that by working with unstructured data, blogs may pose a serious challenge for those who attempt to perform sentiment analysis of blog postings and their commentaries [6]. Pharmaceutical companies, for example, eager to gather market intelligence data about customer satisfaction with new drugs will turn to text analytic programs that measure the emotional content of blogs, known as their “tonal quality.” Such programs, however, may be compromised by the same vagaries of natural language that can compromise speech systems.

At present, speech analytic programs that mine natural language dialog for signs of distress, frustration, anger, and a host of other human emotions, are still of very limited effectiveness. Since they simply match the speaker’s natural language input against the program’s list of keywords (that is, relying solely on word-spotting to process a speaker’s input), they cannot process speech as it actually occurs; if a speaker fails to use the word(s) found in the speech application’s vocabulary, the system is “stumped,” and a poor statistical word match (or no match) is given for the natural language entry. These same limitations apply to programs that look for changes

in prosody, such as a sudden elevation of inflection and/or pitch, insofar as prosodic patterns, like lexical entries, vary across populations of speakers.

Some of the more recent mining programs claim to control for the variations in speaker prosody simply by taking into account that among angry callers, for instance, there are some who may actually lower their tone/pitch and speak more slowly, instead of raising their voices and accelerating their rate of speech [7]. However, such methods can be misleading; they can reduce interactive dialog to simplistic metrics that are inadequate to the needs of speech analysis.

Such shortcomings in speech analytic programs can have significant consequences: when specific keywords or changes in prosody are not found, the speech system can readily overlook important affective data that are critical to accurate assessments of the attitude and affect of the speaker; or, conversely, when keywords and prosodic changes give the specious appearance of caller frustration/anger, mining results can be seriously skewed. Furthermore, one must not underestimate the ripple effect of such limitations on language translation programs, which require in addition to a high word recognition accuracy rate the correct reading of speakers' emotions and intentions to perform proper translation of natural language dialog.

SPA offers a new methodological approach to help surmount these inherent difficulties in speech (and text) analytic systems. Relying on the sequence package in its entirety as the primary unit of analysis, SPA detects meaning and tonal quality in the timing, frequency and arrangement of certain components of dialog or blogs. SPA, which parses a conversation or a blog for its relevant sequence packages, extracts data that are then included in the speech engine's (or text analytic) output stream, contributing to a better understanding of the emotional state of the speaker or blogger. Given that dialog and blogs more often consist of a *blend* of sequences folding into one another than a string of isolated keywords or phrases, one can plausibly argue that speech applications and text analytic mining programs equipped with SPA can better accommodate how people really talk. In particular, SPA-driven programs can better interpret convoluted, repetitive, and elliptical speech – often occurring when a speaker or blogger is seeking professional guidance for a highly distressing medical problem or customer care help-line information to resolve a frustrating consumer-related problem – than conventional speech systems (or text analytic programs) that are limited to a well-defined set of key words and phrases.

In this paper, the author presents one case illustration for each of the three subject domains – medical interviews, customer care contact centers, and consumer-related blogs – to which enhanced speech (and text) analytic programs can be applied. Since there are as yet no hard statistical data to prove the effectiveness of SPA over conventional mining programs (this new NLU method is still in its proof of concept testing phase), the purpose of this paper will be to provide a qualitative

analysis of how SPA may enhance speech (and text) analytic programs by capturing complex affective data that commonly elude such systems.

## II. METHODOLOGY

The way SPA adjusts to speech that is less than “perfect” is to offer a set of algorithms that can work with, rather than be hindered by, the ambiguities, ellipses, idioms, metaphors, colloquialisms, and the many other facets of natural language dialog. Ironically, SPA mines conversations and blogs to find the very sort of dialog and blog data that would have been discarded (or simply ignored) by most speech and text analytic systems as meaningless diatribes, unwieldy talk, or talk that is far too amorphous to grasp. And while some of these discarded data (such as the heightened occurrence of inter-sentential clausal connectives, multiple use of anaphors, idioms and metaphors, or deviations from normal variations in inter- and intra-utterance spacing) might appear relatively unimportant to speech and text analytic programs, these data can be very significant in properly interpreting the emotional content found in natural language dialog and blogs.

Using SPA, the author has designed a BNF (Backus-Naur Form) table consisting of 60 (sixty) Sequence Packages – a typology of parsing structures representing the *semantic* aspects of communication – that capture the affective data found in natural speech and blogs. The parsing structures contained in each Sequence Package consist of a set of non-terminals – context-free grammatical units and their related prosodic features – for which there is a corresponding list of *interchangeable* terminals: words, phrases, or a whole utterance.

The SPA-based BNF table that is used to capture speakers' affect and other semantic aspects of communication – like the BNF tables that are widely used to denote *syntactic* parts of natural language grammars – consists of an elaborate formulation of parsing structures, providing for the incremental design of complex grammatical structures from their more elemental units. In this way, many of the subtleties, convolutions and complexities of human emotion can be more effectively represented by such multi-tiered grammatical structures. A “very angry complaint,” for example, would be depicted on the SPA-designed BNF table as the natural accretion of more elemental parsing features, such as assertions, exaggerations and declarations, so as to effectively notate these semantic aspects of communication – aspects that have all too often eluded conventional spoken language systems.

It is no easy task to formally map out the conversational sequence patterns of natural language dialog and blogs that reflect such elusive, sometimes confounding, human emotions. To do this, SPA draws from the field of conversation analysis as its methodological basis. What conversation analysis provides is a rigorous, empirically-based method of recording and transcribing 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 [8].

By breaking down natural language communication into its elemental form of conversational sequences and speaking turns within those sequences, rather than isolated sentences or utterances, conversation analysts have been able in their 35 years of study of interactive dialog to identify and describe how participants in a dialog systematically accomplish their interactive work, while they are continually engaged in the process of making sense of the ongoing social activity. This is done by examining how speakers demonstrate, through the design of their speaking turns, their understanding and interpretation of each other's social actions, including the wide range of emotions embedded within those actions, such as a speaker's noticeable failure to answer a question directed at him or her, which in certain instances may indicate annoyance or irritation with the other speaker, rather than failure to hear the question.

Reduced to algorithms, many sequence packages are naturally transferable from one contextual domain to another, which means that many of the same sequence package structures found in the conversations of doctors and patients or in patient's blogs, also appear in call center dialog between distressed customers and call center agents. In addition, by focusing on social action, rather than on grammatical discourse structure solely, this new NLU method for mining conversations and blogs can potentially be applied to a myriad of other languages, including Arabic and Farsi, because "all forms of interactive dialog, regardless of their underlying grammatical discourse structures, are ultimately defined by their *social architecture*" [9].

### III. DESIGN

There are two ways that an SPA-driven speech or text analytic mining program can work. First, it can serve as an "add on" layer for conventional data mining programs, including those built on vector-based models, which assign n-grams and bi-grams and hold spaces in between words and word phrases accordingly. If SPA functions as an "add on" layer, the "global weighting" to be applied for the next layer of analysis need no longer be limited to content words or their term roots; rather, it can now also encompass sequence package material. To accomplish this, SPA uses Statistical Language Modeling (SLM) – the standardized method for matching speech input to the speech application vocabularies – but instead of generating candidate words and word phrases for the speech input, SPA generates *candidate* sequence packages. Thus, using the same method of weighting possibilities used for candidate words and word phrases, SPA detects the range of possible sequence packages present at each stage of the conversational sequence, the totality of which makes up the dialog [10].

As an "add on" layer, SPA can take the output of a speech engine and provide a deeper level of analysis of the patient's dialog or blog message by interpolating sequence package information into the output stream. By marking sequence package boundaries and specifying

package properties, the SPA-enhanced mining program gives the software downstream the contextual indicia – the precise location points in the flow of interactive dialog, which signify the different conversational activities and phases of the dialog – needed to interpret the rest of the data stream reliably.

Given that much of blog material, likewise, consists of different phases, including but not limited to narrative complaints, digressions, and complaint resolution, it is essential that a text analytic program, in addition to taking note of the descriptors used by the blogger, provide contextual indicia for analyzing the tonal quality of blogs. For example, since the more strident tonal features are most likely to occur in the digressive phase of the blog, as opposed to the narrative complaint or the complaint resolution phase, a program designed to identify the different phases of blogs will likely give less credence to product criticism occurring in the digressive phase, where diatribes and rants are to be expected. Thus, failure to isolate the various phases of the blog can quite plausibly skew sentiment analysis, by ascribing undeserved importance to the descriptors found in the digressive portion of the blog [11].

Second, SPA might be used as a wholly integrated system rather than as an "add on" layer to conventional speech and text analytic programs. In such a case, such programs would use sequence package grammars, represented by the specially designed BNF table of parsing structures, rather than words and word phrases as their starting point. Such a use would allow the building of an entire vocabulary, by methodically uncovering the keywords characteristically embedded within these sequence package templates, without necessarily having to have an *a priori* knowledge of the words and word phrases in the speech application.

But regardless of whether SPA is built into a system as an "add on" layer of intelligence or in the alternative as a wholly integrated system, it can be argued that SPA, for the most part, can enhance the scalability of data mining and text programs. This is so because SPA can help to streamline the corpus of data required to build a statistical language model, by focusing on commonly occurring sequence packages that are generic to a large population of speakers and bloggers, and thereby eliminate the need to construct elaborate speech application vocabularies, in anticipation of every possible word to be used by a speaker or blogger.

### IV. DOCTOR-PATIENT DIALOG

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, relocation, marriage/divorce, 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 might 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.

Here is an illustration of doctor-patient dialog, derived from tape-recorded interviews of patients at a New York teaching hospital. In this example, which is reflective of a common pattern, the patient offers 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 patient, despite ample opportunity to give this vital history data when the doctor was looking for it, delayed mention of these important facts – presumably fearing that such facts would portend a serious diagnosis for her. In this example, vital family medical history about bone cancer was withheld when the physician queried the patient about family history, causing the doctor (questioned by this author) to either minimize or ignore the importance of this history datum when given later on in the interview. It is beyond question that such delayed patient 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 [12], 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.

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 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, an SPA-driven mining program could retrieve from the recording of this interview this otherwise "lost" data, and its strong emotional implications for the patient.

Here's how. First, SPA would detect a high usage of narrative phrases in close proximity to one another by matching the parsing structures consistent with "troubles-related narration" against the speech input derived from the recording of the doctor patient interview. The arrangement of parsing structures forms the following sequence package, which is divided into four parts. The narrative predicates appear here in ***bold italics***:

1) a short, condensed and somewhat nonspecific concern preceded by a narrative phrase

***I become terribly worried about my pain***

2) an expansion of the concern, citing the troublesome datum ("bone cancer"), which is embedded within two narrative predicates

***which reminds me*** of the arthritic pain that my sister had

***which turned out to be*** bone cancer

3) a recycling of the nonspecific concern preceded by a narrative phrase

***so I worry whenever*** I have any pain

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

*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, in combination with an elevated use of anaphors, the SPA-driven mining program would activate 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 (arthritic pain and bone cancer) embedded within the two narrative predicates ("**which reminds me of...which turned out to be...**") appearing in the *second* part of this four-part sequence package. In addition, such grammars would uncover those specific content words ("worried") and their term roots ("worry") – associated with a high emotive index – 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.

#### V. CUSTOMER-CARE CONTACT CENTER CALLS

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. All in all, retention risk rates (the rates at which customers desert an enterprise in favor of a competitor) might be lowered if the early warning signs of customer dissatisfaction could

be properly detected in recordings of customer care help-line calls between callers and customer service representatives.

The call center industry has long recognized that emotion detection presents one of the greatest challenges for speech analytic programs [13]. A pivotal issue for a call center is knowing when to "*escalate*" a customer complaint call by immediately transferring the call to a supervisor. Failure to recognize early signs of caller distress and to properly route the call to a supervisory agent may seriously compromise customer loyalty. Yet, as long as call center mining programs merely calculate the occurrence of specific keywords/phrases such as the number of times the caller requests a "supervisor" or asks for the "cancellation of an account," those callers whose word choices do not conform to this set of expected keywords will most likely go unnoticed.

Conversely, callers with a propensity to overuse certain words/word phrases, such as "I'd like to speak to a supervisor," even when they are not particularly frustrated or vexed, can easily mislead an emotion detection mining program and consequently generate erroneous scores on caller frustration indices.

Such false positive and false negative readings of callers' emotions in call center recordings underscore the need for a new form of speech analytics that can capture the more generic components of dialog *to* detect the wide range of caller emotion with far greater predictability and accuracy. SPA accomplishes this by relying on the entire sequence package – representing the more generic components of dialog – rather than on individual words or phrases that may vary from speaker to speaker. The example below is drawn from a software help-line for some of the earlier versions of the Microsoft Windows program [14].

**Caller:** "Absolutely unbelievable! What is your name?"

**Agent:** "Mr. Smith."

**Caller:** "Well! I intend to take this much further; this is just absolutely ridiculous!"

In this illustration, while there are no standard "catch" phrases or keywords to signify an irate caller, the caller's exasperation with the customer service agent can nevertheless be detected by an SPA-enhanced mining program, which begins by identifying the sequence package boundaries – and the specific properties found within the parameters of such boundaries – for anger and/or frustration. That is, by looking for specific sequence packages – an organized arrangement of specific (context-free) parsing structures, which more *broadly* reflect the semantic features of communication than a limited set of keywords – the SPA-enhanced program can better detect the caller's true emotional state.

Here's how. First, SPA would detect signs of caller anger/frustration in unusual features that show

aberrations in the dialog. In this example, across several speaking turns, SPA could identify a set of “exaggerative qualifiers.” The first part of the set is the opening exaggerative qualifier (“Absolutely unbelievable!”); the second part of the set is the exaggerative qualifier (“This is just absolutely ridiculous!”) that closes the set. Such a set of qualifiers is *particularly* noticeable when dialog has been progressing up to this point more or less routinely, as in this example.

Second, SPA would look within this set of exaggerative qualifiers for the occurrence of any aberrations in the natural progress of dialog. In this example, an SPA-enhanced mining program would have detected a sequential “non sequitur,” an utterance that is in a sequentially “displaced” position within the dialog: an interposition of a question between this set of exaggerative qualifiers: “What is your name?”

Since requests for speaker identification are generally made at the beginning of a conversation rather than halfway through a call, a caller demonstrates his/her attendance to the social organization of dialog by prefacing a non sequitur request for speaker identity with an apology or excuse (“Pardon me, but I didn’t get your name when you first came on the call” or “What is your name again, I seem to have forgotten?”). It is the noticeable absence of such apologies/excuses preceding the caller’s request for the agent’s name midway in this dialog that raises the caller’s request for the agent’s name from that of a simple inquiry to that of a confrontational argument.

Third, following the non sequitur request for identity information, SPA would look for the occurrence of a declarative assertion, or a threat to take action (“Well! I intend to take this much further...”) which appears prior to the second part of the exaggerative qualifier set, closing off the sequence package for caller anger/frustration.

By mapping out sequence package boundaries, as in this illustration, a speech analytic program can take what ordinarily might appear to be disorganized dialog or simply “shoot-from-the-hip” talk – common when a caller is preoccupied with “venting” his frustration at the contact care center agent – and find the indicia of caller emotion buried in such unwieldy discourse.

## VI. CONSUMER HEALTH-RELATED BLOGS

Consumers’ blog postings and comments on health-related topics present a goldmine of data for medical researchers, practitioners, and pharmacologists. Nowhere can one find such immediate feedback on new drugs and medical devices than in the blogosphere. Ironically, those who eschew surveys, refusing to answer questionnaires handed to them at their doctor’s office or received in the mail, will freely post their personal data on blog sites without giving it a second thought.

The same methodological approach used to analyze natural speech may be applied to blogs. While this article focuses on consumer health-related blogs, SPA may be applied to many other blog domains: consumer car purchases, home renovations and electoral politics, to

name just a few. Drawing from the BNF table of parsing structures that represent the semantic aspects of communication, and their attendant affective features, an SPA-driven text analytic program begins by marking off sequence packages that indicate the different phases of the blog, such as narrative complaint, digression, and complaint resolution. To assess the true tonal quality of the blog, a score, based on the frequency and intensity of adjectival descriptors, is generated for each phase respectively, *before* calculating the final tonal quality score.

Below is an example of a blog posting [15]. For analytic purposes, the blog (originally a continuous text) has been broken into three distinct phases:

### A. Narrative Complaint

“I was prescribed Ambien a couple of years ago. I shared my concerns with my doctor... And I started having hallucinations... I refused to take Ambien any longer and started on another prescription for sleep...”

### B. Digression

“Here is another example of the FDA knowing that a drug is this dangerous, but because the pharmaceutical companies can make tons of money off of it, these side effects are kept silent. The prescriptions we are taking are slowly killing us... Each prescription we take causes side effects and other diseases. The sick will just keep getting sicker...”

### C. Complaint Resolution

“Unfortunately, I have not been able to find natural alternatives yet that help me sleep effectively, but I haven’t had any prescriptions that do that either. This might be one area where using both together might be my only recourse.”

In this example, the blogger, who had been taking Ambien (a frequently prescribed sleep-inducing agent), used strong, inflammatory language, filled with adjectival descriptors that connoted anger and frustration. However, a closer look reveals that the blogger reserved her strong emotive descriptors (which seem to allege a sinister plot on the part of the pharmaceutical companies in conjunction with the FDA) mainly for the “digression” phase of the blog, rather than the narrative complaint or complaint resolution phase. Since conventional text analytic programs commonly add up the total number of descriptors in the blog, *without* first marking off the sequence packages that differentiate, for example, between digressions, narrative complaints, and complaint resolution, sentiment analysis can be distorted by the excessive use of strong descriptors in the digression phase [16].

In this example, by contrast, SPA would minimize (or even discount) the *potency* of the language used in the digression phase, focusing more accurately on the temperate and reasoned tonal quality of the narrative complaint and the complaint resolution (which shows the blogger to be balancing the benefits and risks of various kinds of sleep aids). As a result, the total score given for

the blog's tonal quality would be much closer to the true affective state of the blogger. In so doing, SPA enables text analytic programs to perform more accurate sentiment analysis, which in turn produces more reliable market intelligence data.

## VII. CONCLUSION

The rapid advances in audio (and text) mining technology – where large amounts of audio recordings and unstructured text, as in blogs, are filtered through speech and text analytic programs – are characteristic of an industry best defined by “Moore’s Law,” the principle of exponential growth in computer processing power over the past 40 years. As shown in this article, audio and text mining technology has advanced so rapidly that extracting complex emotional data from calls and blogs is no longer a goal of the elusive future. Rather, mining for the subtleties of human emotions as expressed in natural language dialog – a task once considered far too complicated for speech systems – is now the next frontier of speech (and text) analytics, whether for call centers, medical interviews, or consumer-generated blog postings.

In this paper, the author has shown how a new natural language understanding method, known as Sequence Package Analysis, provides the grammatical underpinning for simulating human listening and interpretive skills. SPA relies on a BNF table of parsing structures to reflect the complex semantic aspects of communication. As a new AI-based natural language understanding method, SPA offers a set of flexible grammars that can parse small segments or an entire dialog to detect emerging patterns in the conversation, such as a buildup of customer frustration with a call center agent. By detecting such patterns, depicted in the form of sequence packages, the speech system may glean crucial business intelligence data that would have otherwise required costly and time-consuming human analysis.

Clearly much remains to be done. But the author predicts that the capabilities of speech and text mining programs to detect complicated and mercurial human emotions from recordings and consumer blog postings will justify Moore’s Law, advancing exponentially over the next decades. Whether this phenomenon will one day be called “Neustein’s Law” is harder to predict. But SPA is a critical tool in the advancement of the emotion-detection capabilities of speech and text mining programs that are the industry’s next frontier.

## REFERENCES

- [1] A. Neustein, “Using Sequence Package Analysis to Improve Natural Language Understanding,” *International Journal of Speech Technology*, 4(1), 2001, pp. 31-44.
- [2] M. Paprzyki, A. Abraham, and R. Guo, “Data Mining Approach for Analyzing Call Center Performance,” *The 17<sup>th</sup> International Conference on Industrial Engineering Applications of Artificial Intelligence and Expert Systems*. Ottawa, Canada. In *Lecture Notes in Computer Science*. Germany: Springer Verlag, 2004.
- [3] A. Aschenbrenner and S. Miksch, “Blog Mining in a Corporate Environment,” *Vienna University of Technology, Institute of Software Technology and Interactive Systems and Research Studios Austria, Smart Agent Technologies*, Vienna Technical Report, Asgaard-TR, November 2005.
- [4] D. Sifry, “State of the Blogosphere,” *Sifry’s Alerts*, August 7, 2006.
- [5] S. Herring, et al., “Conversations in the Blogosphere: An Analysis ‘From the Bottom Up,’” *Proceedings of the 38<sup>th</sup> Annual Hawaii International Conference on System Sciences*. Los Alamitos: IEEE Computer Society Press, January 3-6 2005.
- [6] “Exploring Attitude and Affect in Text: Theories and Applications,” *AAAI Spring Symposium*. Stanford University, March 22-24, 2004.
- [7] P. Britt, “Advanced Analytics Offer Greater Precision,” *Speech Technology Magazine*, 12 (4), May 2007, pp. 32-35.
- [8] J. M. Atkinson and J. Heritage, Transcript notation. In J.M. Atkinson and J. Heritage (Eds.), *Structures of Social Action: Studies in Conversation Analysis*. Cambridge University Press, Cambridge, 1984, pp. ix-xvi.
- [9] A. Neustein, “Sequence Package Analysis: A New Global Standard for Processing Natural Language Input?” *Globalization Insider*, vol. X111(1, 2), February 18, 2004.
- [10] A. Neustein, “Using Sequence Package Analysis as a New Natural Language Understanding Method for Mining Government Recordings of Terror Suspects,” *Proceedings of the 3rd International Workshop on Natural Language Understanding and Cognitive Science, ICEIS 2006*, Paphos, Cyprus, pp. 101-108, May 23-24, 2006.
- [11] A. Neustein, “Performing Sentiment Analysis of Blogs,” PowerPoint Presentation, *Linguistic Technology Systems*, April 2006.
- [12] R. Frankel, “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, 1990, pp. 231-262.
- [13] J. Markowitz, “Detecting Emotion,” *Speech Technology Magazine*, 11 (1), January/February 2006.
- [14] M. Emmison, “Calling for Help, Charging for Support: Some Features of the Introduction of Payment as a Topic in Calls to a Software Help-Line,” *Symposium on Help-lines*, Aalborg, Denmark, September 8-10, 2000.
- [15] [http://fightingfatigue.typepad.com/my\\_weblog/2006/news\\_on\\_ambien.html](http://fightingfatigue.typepad.com/my_weblog/2006/news_on_ambien.html)
- [16] A. Neustein, “Sequence Package Analysis: A New Natural Language Understanding Method for Intelligent Mining of Recordings of Doctor-Patient Interviews and Health-Related Blogs,” *Proceedings of the Fourth International Conference on Information Technology: New Generations, ITNG 07*, pp. 441-448, April 2-4, 2007.

**Amy Neustein**, Ph.D. is the Founder and CTO of Linguistic Technology Systems ([www.lingtechsys.com](http://www.lingtechsys.com)), a NY-area based think tank for advances in audio (and text) data mining programs for healthcare, consumer relations, and government security. Dr. Neustein has authored and edited academic books and many articles for academic and industrial publications. Her research on data mining was most recently published by the IEEE Computer Society in the proceedings of the Fourth International Conference on Information Technology: New Generations.