

How Much Do We Say? Using Informativeness of Negotiation Text Records for Early Prediction of Negotiation Outcomes

Marina Sokolova · Guy Lapalme

Published online: 21 August 2010
© Springer Science+Business Media B.V. 2010

Abstract Business negotiations represent a form of communication where informativeness, i.e., the amount of provided information, depends on context and situation. In this study, we hypothesize that relations exist between language signals of informativeness and the success or failure of negotiations. We support our hypothesis through linguistic and statistical analysis which acquires language patterns from records of electronic text-based negotiations. Empirical results of machine learning experiments show that the acquired patterns are useful for early prediction of negotiation outcomes.

Keywords Electronic negotiations · Text data mining · Machine learning · Language patterns · Early prediction of success or failure

1 Introduction

We use language to convince, explain, question, bargain and, doing this, establish and reach our goals. The goal-oriented aspect of the language use is even stronger in the context of negotiations, especially when negotiators use only written means to communicate, that is to say, exchange letters, email, text messages. In face-to-face and visual conference negotiations, prosody and body language often play a critical role in conveying attitudes and feelings. In contrast, negotiators who must rely only on texts obtain their information from numeric offers and the language in which these offers are written or accompanied.

M. Sokolova (✉)
Children's Hospital of Eastern Ontario Research Institute, Ottawa, ON, Canada
e-mail: msokolova@ehealthinformation.ca

G. Lapalme
DIRO, Université de Montréal, Montréal, QC, Canada
e-mail: lapalme@iro.umontreal.ca

Table 1 A sample from Inspire negotiation. The data used in this research has 2,557 negotiations in total, 1,427 of them successful, others—failed. The average length of successful negotiations is 718 words, the average length of failed negotiations is 594 words

Buyer	<i>Hi Joe, Im Lisa and I represent Cypress Cycles in this negotiation. After extensive deliberation we have prepared an offer to purchase sprockets and gear assemblies. We think it is a fairly good offer and hope you find it acceptable.</i>
Seller	<i>Hi Lisa, I am Joe, the representative of Itex Manufacturing and I am very delighted to get in touch with you. First of all, thank you very much for the possibility to negotiate with you and your company. Despite your really interesting offer, it is not possible for me and my company to accept it under all circumstances. For that reason I would like to make the following proposal to you. I am very interested in what you are thinking about, so I am looking forward to hearing from you. Bye, Joe.</i>

In this work, we perform informativeness analysis of language used in *text-based electronic negotiations*, i.e., negotiations conducted by text messages and numerical offers sent through electronic means. We use language pragmatics to define linguistic expressions of the message informativeness, and then apply learning software to identify the expressions that are indicative of the final outcome of the negotiation—success or failure. In this paper we focus on the negotiation as an ongoing process. We analyze the linguistic features of messages exchanged in the course of the negotiation, to determine whether the first or the last part of negotiations is more important for the prediction of the outcome.

We want our empirical results to be general and applicable to a wide range of negotiation settings. We seek negotiation texts where language is not confined within the scope of this negotiation, but is general enough to be used in other negotiation settings. An important indicator of a possible result generalization is the text vocabulary growth. If a gradual growth of the text sample size causes vocabulary to grow, then language results obtained on such texts can be used outside the particular negotiation, as stated by Oakes (2003). Hence, we look for negotiation records whose vocabulary grows with the increasing text sample. Inspire negotiations supply us with such texts. Those are asynchronous text-based electronic negotiations had the largest available textual data (Kersten and Zhang 2003). Statistical and syntactic language analysis of the Inspire negotiation text records has shown that the vocabulary monotonically grows with the text size (Sokolova 2006). Table 1 presents a text sample of one negotiation.

Some preliminary results of the current work were reported in Sokolova and Lapalme (2007).

2 Background Review

Communication, through a variety of forms, conveys messages sent by a speaker and received by a hearer. These messages can be complex and subtly expressed and made up from what is said and what is implied, according to Leech and Svartvik (2002). Success of communication depends on the speaker's ability to produce a message and on the hearer's ability to understand it. Pragmatics, the study of language use, accepts

that to be able to infer the meaning of a speaker's message, the hearer expects that the message should satisfy standards of the Grice Maxims, introduced by [Grice \(1989\)](#):

- Quantity (informativeness),
- Quality (truthfulness),
- Relation (relevance),
- Manner (clarity).

Not all communications do so, sometimes a hidden context interferes with the correct understanding of a message. In the process of negotiations, paying more attention to the wording of offers allows for better understanding of the situation and improves prediction of the future actions. This makes linguistics an important tool in negotiation studies. Its role is critical for studies of text-based negotiations, especially computer-mediated negotiations in which application of non-linguistic means can be restricted. To obtain valid results, it is important to correctly project reasoning onto text and language features, as suggested by [Mahmoudi et al. \(2008\)](#).

As a special type of communication between people, negotiation is a dynamic process that is multi-dimensional, irreversible, and purposeful, according to [Rolloff and Putnam \(1992\)](#). The way negotiators interact depends on many factors, such as

means, face-to-face meeting, email;

topic of discussion, business, personal;

communication mode, synchronous or asynchronous;

interaction mode, one-to-one, one-to-many;

speaker-hearer roles, doctor-patient, buyer-seller, presenter-audience.

Currently, learning software applied to study language in negotiations is mostly used in restricted language environments, e.g., agent–customer phone conversations, where an agent follows the call flow pre-defined by his company's policy ([Takeuchi et al. 2007](#)). In planning dialogues ([Chu-Carroll and Carberry 2000](#)), the discourse analysis uses Searle's theory of speech acts to support the fact that language carries much of people's behaviour and emotions. In [Reitter and Moore \(2007\)](#), the authors studied repetitions in task-oriented conversations which can be considered as a part of a negotiation process. They demonstrated that a speaker's short term ability to copy the interlocutor's syntax is autonomous from the success of the task, whereas long-term adaptation varies with such success. We, however, aim to study language expressions in negotiations which represent general, non-restricted language environment, following [Sokolova \(2006\)](#).

In studies of electronic negotiations, comparison with face-to-face negotiations often brings in new insights of the medium influence on the negotiation process. Research on text records of face-to-face negotiations ([Simons 1993](#)) suggests that the language patterns used in the first half of a negotiation predict the negotiation outcome better than those in the second half. The explanation was that in the first phase people establish contact, exchange personal information and engage in general polite conversation, creating a foundation of trust between partners. No numerical data, however, supported this diagnosis, and there was no distinction between the prediction of successful and unsuccessful outcomes. However, [Sokolova et al. \(2008\)](#) have reported different empirical results. The results show that when e-negotiation

texts are represented by language signals of negotiation strategies, the classification of the second parts is more accurate with respect to the outcome than the classification of the first parts. Alongside with the communication medium, communication mode also plays an important role. According to Koeszegi et al. (2007), asynchronous electronic negotiations usually do not have a sequential-stage model of behaviour, which is common in face-to-face negotiations, (Adair and Brett 2005) and synchronous electronic negotiations. Here is an example of behavioural phases in synchronous negotiations, presented in Koeszegi et al. (2007):

Perform Relational Positioning → Identify the Problem
 → *Generate Solutions → Reach Agreement*

Unexpected turns and moves—typical of human behaviour—make prediction of the negotiation outcome difficult. In case of asynchronous electronic negotiation, the absence of the usual negotiation phase structure further complicates the outcome prediction. Although analyzing the negotiation process from many different points, the works cited above do not consider the information aspect of negotiation language. In contrast, we concentrate on the informativeness of language used in negotiations. We analyze information exhibited by word categories such as adjectives and adverbs that we consider as indicators of text informativeness. This representation is then used in statistical and machine learning experiments for establishing relations between informativeness and success or failure of negotiations.

3 Assessing Text Informativeness

The amount of information given by a message is called the *informativeness* of a message. For a structured or guided message, the informativeness of a message correlates with the speaker's position revealed by the message or with the speaker's opinion delivered by it (Loken 2006). As suggested in Kamakura et al. (2006), the relation between informativeness of free text message and the corresponding opinion and position is more subtle. In this article, we work with free-form texts of text-based electronic negotiations. We study the relation between informativeness of negotiators' communication and negotiation outcomes.

Assessing message informativeness is a complicated task due to the fact that the conveyed information consists of two related parts: said information and its implications (Sperber and Wilson 2006). *What is said* defines the quality and quantity of possible inferences and the message information, which is perceived within a given context and established linguistic rules. The combination of the three factors—information, linguistic meanings, context—allows the hearer to infer and recognize *what is communicated*. As a result, informativeness is determined by *what is said* and *what is inferred* and is understood within a communication context. Pragmatics, the study of language use, suggests that to be able to infer the meaning of the speaker's message, the hearer expects the message to be as informative as necessary for the situation at hand, trustworthy, relevant, clear and brief as stated by Grice (1989). However, informativeness of messages varies. For example, in Table 2 message (a) implies that

Table 2 Similar messages with different informativeness on the educational reform; (a), borrowed from Carston (1998), is more informative than (b)

The educational reform has been successful.

(a) *Many schools are doing better than before.*

(b) *Many schools are doing very well.*

Table 3 Examples of words in word categories signaling informativeness of a message. The resulting lists include 65 words for degrees, 34 for scalars and 40 for comparables.

Category	Examples
Degrees	<i>almost, always, only, just, ...</i>
Scalars	<i>and, one, all, necessarily, ...</i>
Comparatives	<i>among, greatest, identical, ...</i>

The words were taken from Roget's Interactive Thesaurus (2006)

although not all schools are doing better than before the reform, more than just few improved their situation after the reform had started. On the other hand, message (b) does not imply anything about the reform impact, because we do not know whether the schools were doing excellent or poor work before the reform. As the result, (a) provides more information than (b) about the dynamics of the area schools.

Different information types cause different inferences and vary in their contribution to the message informativeness (Sperber and Wilson 2006). The types are represented by linguistic expressions, that can be analyzed by estimating the strength of the used words. The guiding principle for linguistic analysis is that a true stronger language statement makes all comparable weaker statements true (Carston 1998). Thus, comparative comments and estimations are more specific and precise and bear more information than declarative ones. Words used in comparative comments and estimations (Table 3) contribute to informativeness of a message. Informativeness is estimated with respect to the context and situation of communication, e.g., its goal, the type of participants' interaction, communication means and rules. Our working assumption is that negotiation results are connected and related to different types of the information.

4 Text-Based Electronic Negotiations

Convenience of email, instant messaging resulted in a fast-growing number of participants in text-based electronic negotiations, i.e., negotiations conducted through electronic means by exchange of texts which may include numerical offers. People negotiate through email or negotiation-support systems in legal and economic settings as well as in research and training. The use of electronic means changes the way people communicate during negotiations. In Tables 4 and 5, we compare sample transcripts of a face-to-face negotiation from Marriot (1995) and an electronic negotiation from Sokolova (2006). The transcripts illustrate how much information during a face-to-face meeting can be gained from non-verbal body language (gestures, movement and so on) and language characteristics (length of pauses, tone of voice and so on).

Table 4 Exempts from transcripts of face-to-face bilateral business negotiations

Face-to-face negotiations		
Roles	Language exchange	Additional cues
Buyer	<i>It eh what the container eh quan- quantity of each block?</i>	
Seller	<i>Two kilos.</i>	<i>(J outstretches his two arms to indicate a block)</i>
Buyer	<i>Right.</i>	<i>(After clarifying details relating to the size of the product, the Japanese businessman writes a note in his notebook)</i>
	<i>And eh ah so you you don't have any propriety of eh the license of the- or another patent, but you have ah know-how to make this</i>	<i>(Lays his pen down)(10 second pause)</i>
Seller	<i>Yeah, yeah, there's many people who have tried to make it...</i>	<i>(Glances at the first page of his notes)</i>

The right column shows additional information extracted from communications

We also see that language plays a bigger role in text-based electronic negotiations, expressing negotiator's bargaining, formal introductions and closures moves, and allowing extra activities as socializing and exchange of personal information. This bigger role gives us an opportunity to seek additional cues embedded in the language exchange (Table 5).

The current study concentrates on the *pragmatics* of communications, with a focus on comparative comments and estimations that are mapped to degree, scalar and comparative word categories. These word categories include adjectives, adverbs, conjunctions (degrees and comparatives) and cardinal numbers, determiners, cognition verbs, conjunctions (scalars) that are used to compare events and objects (Carston 1998). These pragmatic cues substitute for visual information available in face-to-face negotiations. We support our hypothesis by experiments on data of electronic negotiations.

5 Text Data of Inspire Negotiations

The largest data set gathered in e-negotiation comes from *Inspire*, a public-domain research and teaching tool mostly used in college and university programs in numerous countries (Kersten and Zhang 2003; Hine et al. 2009). It allows its users to conduct negotiations over the Web, gives access to on-line manuals, provides automatic evaluation of the negotiation process, and keeps a log of each negotiation. No restrictions are imposed on users.

Previously, *Inspire* data has been used in data mining and machine learning applications. In Kersten and Zhang (2003), the authors analyzed outcomes of negotiations conducted using *Inspire*. They applied several data mining algorithms to the history of the exchange of formal offers. Among their findings the following results on the behaviour of e-negotiators are very interesting: if offer exchanges are made during

Table 5 Exempts from record of electronic bilateral business negotiations held through Inspire

Electronic negotiations		
Roles	Language exchange	Additional cues
Seller	<p><i>Hi Anles, I have just sent a counter-offer to you. It wasnt such easy, as I thought cause it seemed I made my ratings wrong *g*. Well, now I already asked you, where you are from, cause I did not know that I would have the opportunity to contact you again. I am from Germany. Then, good luck with my offer, I am waiting for your answer. Bye Claudi</i></p>	<p>Just – degree, as – comparative, thought – scalar</p>
Buyer	<p><i>Hi claudi, thank you very much for your offer. I think, the price is acceptable. I totally agree with you. Having informed at a trade fair in Frankfurt/Germany about metal components and comparing some prices and offers from other suppliers all around the world, I came to the conclusion that your offer is the best. It was a pleasure doing business with you. I'll give you a ring this week for more details. Best regards anles</i></p>	<p>Think – scalar, totally – degree, and – scalar, some – scalar, all – scalar, best – comparative</p>

Language signals of comparativeness are shown in *bold*. The right column shows additional information extracted from communications

the early stages of the negotiation, there is a higher possibility of reaching an agreement; offers sent during the last day before the deadline reduce the probability of achieving an agreement. In Sokolova and Szpakowicz (2006), the authors studied tactical moves and influence strategies of negotiators by application of popular machine learning techniques. They analyzed language patterns corresponding to commands, requests, advices, prohibitions, etc. The extracted language patterns have been used to represent negotiations in a set of machine learning experiments. Empirical results obtained on Inspire data showed that language patterns provide better classification of negotiation outcomes than the 500 most frequent words, including stop words, e.g., *the, of, in*, although these words are accountable for >50% of the negotiation records. Nastase and Shirabad (2007) studied the influence of sentiments and gender on negotiation outcomes. The authors used DECISION TREE to determine negotiation characteristics which provide better accuracy of success and failure classification. Their results showed that knowing genders of both negotiators is a good predictor of success. There was no difference between the same or different genders. When negotiators' sentiments were, only when one negotiator was a male and another gender was unknown, the negative sentiment was picked by DECISION TREE as an important feature. The positive sentiment did not appear as an important feature. Negotiators' positive and negative emotions were studied through the word count in Hine et al. (2009). Researchers applied LINGUISTIC INQUIRY AND WORD COUNT by Pennebaker et al. (2001) to assess positive or negative emotions, agreeable or negative language and in-group or

out-of-group perception of negotiators. The obtained results showed that successful e-negotiations are characterized by proportionally more positive emotion and more agreeable and less negative language, if compared with unsuccessful e-negotiations. However, the results on negative emotions were inconclusive, they were not significantly associated with e-negotiation success; similarly, no significant association was found between e-negotiation success and group involvement indicators.

The *Inspire* text data available to us consists of the transcripts of 2,557 negotiations, 1,427 of them successful (for samples of the data refer to Tables 1 and 5). One person can participate in only one negotiation. For the available data, the number of data contributors is over 5,000. We work with raw, unedited data that contain 1,514,623 words (tokens) and 27,055 distinct words (types). Negotiation is bilateral, between a buyer and a seller of bicycle parts, with four issues (price, delivery time, payment time, return conditions), each with only a few fixed numerical values. Negotiators exchange formal offers (tables with numerical values) and may send free form messages. Exchange of text messages is optional. They either accompany offers or are exchanged between offers. A negotiation, lasting up to 3 weeks, succeeds if a virtual purchase took place within the designated time, and fails otherwise.

Negotiations mediated by *Inspire* provide us with rich data. First, the negotiations are long enough to allow the participants to develop and apply their strategies. The longer an e-negotiation lasts, the more complex the structure of the e-negotiation process becomes. Simpler e-negotiation may involve exchange of well-structured business documents (pre-defined contracts, retail transactions). A complex e-negotiation process comprises numerous offers and counter-offers and has a high degree of uncertainty. Next, the number of participants—more than 5,000—guarantees that the corpus analysis results are not biased by the personal specifics and that they show general trends exhibited by groups of negotiators.

We analyze *Inspire* text records following the quantitative corpus analysis by Oakes (2003); other measures of lexical diversity and their applications are discussed in Yu (2010). We employ N -gram models which are arguably the most widely used models for the language analysis purposes. We construct N -gram models, where $N = 1, 2, 3, 4$, by computing

$$P(w_k | w_1^{k-1}) \approx P(w_k | w_{k-N}^{k-1}) \quad (1)$$

where $P(w_k | w_1^{k-1})$ is the probability of the word w_k appearing after the sequence of words $w_1 \dots w_{k-1}$. Obtaining N -gram frequencies is a necessary and important step in understanding communication data—for each corpus originating from a specific genre or source, the N -gram frequency distribution is one of the essential characteristics. Here we provide the synopsis of the main ideas and their realization by the statistical analysis:

1. To find how degree, comparative and scalar words are used, we first search for the most frequently used words of *these categories*. We seek the words that are representative for negotiation data in general and across its possible subsets (e.g., texts of successful negotiations, messages sent by buyers). In order to find such

- words, we analyze the complete text data and texts of its four subsets, i.e., sent by sellers in successful negotiations, sellers in unsuccessful negotiations, buyers in successful negotiations and buyers in unsuccessful negotiations.
2. We build the unigram model of the data and look for words occurring more than 10 times in each of the subsets. This gives 67 degree, comparative and scalar words. The five most frequent words are *and*, *better*, *only*, *more*, *think*, *than*, listed in alphabetical order.
 3. We then build a bigram model of the data, remove the bigrams occurring less than 4 times and the so-called stop bigrams, i.e., bigrams containing determiners, articles and prepositions; for example, *of the*, *if a* were removed. From the remaining set we extract bigrams containing the 67 designated words. Examples of the most frequent bigrams are *and I*, *better price*, *can only*, *be more*.
 4. To obtain more information on the use of the words, we build the trigram model of the data and the four subsets and extract trigrams containing the informativeness words. Examples of the most frequent trigrams are *and I am*, *your offer and*, *the only way*, *know more about*.

Although there are no obvious trends in the structure of the trigrams, a common use of personal pronouns has emerged as a pattern, which could be explained by a correspondence between bilateral negotiations and dialogue. In the next section, we use these results to represent negotiation data in machine learning experiments.

6 The Outline of Learning Experiments

To find out whether our hypothesis on relations between informativeness of negotiation records and outcomes holds on a bigger scale, we employ Statistical Natural Language Processing and Machine Learning techniques (Manning and Schütze 2003; Witten and Frank 2005). The use of statistical and learning techniques allows to test our assumption on a larger amount of data.

First, we have employed statistical analysis to find language patterns that are characteristic to negotiators' communication. Generalization of the most frequent N -grams helps us to find patterns that correspond to the use of the informativeness words. We use these findings to represent negotiation texts in machine learning experiments.

Next, machine learning experiments provide an opportunity to determine what learning model better explains dependencies between informativeness and negotiation outcomes. We seek dependencies with success/failure negotiation outcomes thus perform *qualitative* analysis of the relations. For each negotiation record, we consider that its success or failure corresponds to the negotiation outcome.

We hypothesize that informativeness relates to these outcomes. To prove this, we use *supervised learning*, i.e., when a learning algorithm \mathcal{A} has an access to the example labels. It constructs a function on training data, a set of input and output pairs (\mathbf{x}, y) where \mathbf{x} represents a negotiation text through the informativeness language signals and y is the negotiation outcome. Then the algorithm uses this function to predict outcomes on testing data of previously unseen examples. Learning negotiation success or failure thus defines a *binary classification* problem that we solve with a variety of learning algorithms:

- decision-based algorithms, DECISION TREE(DT) and ALTERNATING DECISION TREE (ADT);
DT estimates $InformationGain(IG(a, y))$ of the attributes to discriminate between classes:

$$IG(a, y) = H(t) - H(t|a) = H(t) - (p_L H(t_L) + p_R H(t_R)) \tag{2}$$

where $IG(a, y)$ is used for class label in binary classification problem, a is the splitting attribute, y is the value of a , t is the distribution of a , p_L, p_R, t_L, t_R are proportions of elements and the distributions of left and right nodes, respectively. ADT alternates prediction estimates, which are generated by boosting of the first layer of DT and $InformationGain$ of the attributes to split data. Both the decision-based algorithms output models for analytical analysis.

- a learning algorithm NEAREST NEIGHBOR(KNN) that evaluates the class label of an entry based on the labels of entries closest to it; closeness is evaluated by calculating distance between representations;For the input x it calculates the output $\hat{Y}(x)$ by averaging the labels of k closest points x_i in the training set:

$$\hat{Y}(x) = \frac{1}{k} \sum_{x_i \in N_k(x)} y_i \tag{3}$$

where $N_k(x)$ is the neighborhood of x and y_i is the label of x_i .

The algorithm’s output shows a level of similarities among data entries;

- kernel-based methods:
RADIAL BASED FUNCTION NETWORKS(RBFN) is a two-layer fully-connected network with an input layer which performs no calculation. The hidden layer learns by applying the unsupervised K -means algorithm. The output layer learns by computing a liner combination of the weighted basic functions.
SUPPORT VECTOR MACHINE(SVM) builds a hyperplane that separates training examples into two classes, with the largest possible separation. The search for the hyperplane is done by solving a constrained optimization problem in dual representation:

$$\mathbf{w} = \sum_j \alpha_j c_j \mathbf{d}_j, \quad \alpha_j \geq 0, \tag{4}$$

where \mathbf{w} represents the hyperplane, $c_j \in \{1, -1\}$ is a class label, α_j are found by solving a dual optimization problem. Support vectors \mathbf{d}_j correspond to $\alpha_j \geq 0$. SVM, known for high accuracy of classification of texts, serve as empirical estimate of the goodness of results.

We use the algorithms’ implementation available from Weka, an open source learning software presented in [Witten and Frank \(2005\)](#).

Next, we evaluate the algorithm’s performance. Quality of classification can be assessed using a confusion matrix, i.e., records of correctly and incorrectly recognized examples for each class. Table 6 reports on binary classification, where tp are true positive, fp – false positive, fn – false negative, and tn – true negative counts.

Table 6 A confusion matrix for binary classification

Class	Classified	
	As <i>pos</i>	As <i>neg</i>
<i>pos</i>	<i>tp</i>	<i>fn</i>
<i>neg</i>	<i>fp</i>	<i>tn</i>

To estimate how the classification algorithms work, we calculate

$$Accuracy = \frac{tp + tn}{tp + fn + fp + tn}, \quad (5)$$

and measures commonly used in text classification (Sebastiani 2002):

$$Precision = \frac{tp}{tp + fp} \quad (6)$$

$$Recall = \frac{tp}{tp + fn} \quad (7)$$

$$Fscore = \frac{(\beta^2 + 1)tp}{(\beta^2 + 1)tp + \beta^2fn + fp} \quad (8)$$

Text classification measures are imbalanced towards classification of positive examples. Hence, we obtain a better picture of the success classification than the one of failure.

Finally, we analyze the models produced by the algorithms which obtained the smallest measure values. These algorithms better model the negotiation texts than other algorithms. Analyzing the models, we find relations between the informativeness language features appearing in them and the negotiation outcomes.

7 Empirical Prediction of the Negotiation Outcomes

Early prediction of upcoming events is an important learning task in many domains. We want to know whether text informativeness provides a reliable prediction of the negotiation outcomes from the first part of negotiations. We say that early prediction is *reliable* if the classification results are statistically close to those achieved on complete negotiations. In these experiments we use the extracted 67 words to represent negotiations. Prediction of success or failure of negotiations aims to find whether a text belongs to one of the two categories of negotiation texts. This is a *classification* learning task. We consider successful negotiations to be the positive class, and unsuccessful negotiations—the negative one.

In the first part of machine learning experiments presented here, the data consists of the texts of the first half of negotiations. This segment is labelled by the outcome of the whole negotiation. In the second part of the experiments, the data consists of the texts of complete negotiations. For each data entry we assign 67 attributes, one

Table 7 The algorithms' best classification results of success or failure of negotiations

Classifiers	First half				Complete			
	Acc	F	Pr	Rec	Acc	F	Pr	Rec
KNN	0.66	0.67	0.73	0.62	0.66	0.68	0.72	0.65
ADT	0.70	0.74	0.71	0.76	0.70	0.75	0.69	0.83
DT	0.68	0.73	0.69	0.77	0.69	0.75	0.69	0.81
RBFN	0.68	0.75	0.68	0.83	0.68	0.75	0.67	0.86
SVM	0.70	0.74	0.70	0.78	0.70	0.76	0.70	0.83

Negotiation texts are represented by degree, scalar and comparative words. The left part presents classification of the first part of negotiations; the right part presents classification of complete negotiations. For each of the data representations, we show the best value of each measure (in bold)

attribute for each of the degree, scalar and comparative distinct words (types). An attribute is binary: it is 1 if the word appears in the entry and 0—otherwise. This data representation captures whether a comparative event happens in a negotiation or not. We do not keep track how many times each word appears in the same negotiation, because this information could be understood in two different ways: either the negotiator is persistent in pursuing certain goals and keeps referring to them several times or her English vocabulary or grammar, or both, are limited. This research avenue opens many possibilities for analysis, that is why we leave it for future work.

We evaluate the algorithm's performance with respect to the results the classifiers obtained on the dominant class (i.e. successful negotiations). Table 7 reports the best results obtained by a thorough search of the algorithm's adjustable parameters. *Accuracy* (*Acc*), *F*, *Pr*, *Rec* are computed by Eqs. 5–8. The results are estimated by ten-fold cross-validation. We calculate *Fscore* for $\beta = 1$. To put the learning approach in perspective, we consider the baseline when all negotiations are classified as positives. Then, for prediction of the negotiation success, *Accuracy* equals 0.55. Corresponding *F* is equal to 0.71. Based on the best results, machine learning predictions are more accurate for all the reported experiments.

The empirical results show the reliability of the early prediction of the negotiation outcomes when negotiation texts are represented by the informativeness signals. For an algorithm, classification of first half and complete negotiations is not statistically significant different (the paired Student's *t*-test (Oakes 2003)). However, conclusion differs for classification of successful and unsuccessful negotiations. Increase of *Recall* shows that informativeness of complete negotiations assists in a better classification of successful negotiations. This holds for the five classifiers. *Precision* has decreased for NEAREST NEIGHBOR and ALTERNATING DECISION TREE and is steady for the other classifiers. This trend implies that informativeness is more important in the first part of unsuccessful negotiations than for complete negotiations.

Another interesting conclusion can be drawn when we look at the parameters of the best classifiers built by the algorithms. It is important to remember that the structure of models built by each algorithm remains the same on the first half and complete negotiations. NEAREST NEIGHBOR classifies unsuccessful negotiations better than the other algorithms. For each data entry, the algorithm needs 15 entries, closest to it with

Table 8 The algorithms' best classification results of success or failure of negotiations

Classifiers	First half				Complete			
	Acc	F	Pr	Rec	Acc	F	Pr	Rec
KNN	0.71	0.74	0.72	0.77	0.69	0.73	0.71	0.76
ADT	0.70	0.75	0.70	0.82	0.70	0.75	0.69	0.83
DT	0.68	0.74	0.69	0.79	0.69	0.75	0.69	0.81
RBFN	0.69	0.75	0.68	0.83	0.69	0.76	0.68	0.86
SVM	0.70	0.76	0.69	0.85	0.71	0.78	0.69	0.89

Negotiation texts are represented by personal pronouns, degree, scalar and comparative words. The left part presents classification of the first part of negotiations; the right part presents classification of complete negotiations. For each data representation, the best measure values are shown in bold

respect to the Euclidean metric, to find its class label. This holds for the first half of negotiations and a complete negotiation. This implies a *high level of similarity* among occurrences of the informativeness signals in unsuccessful negotiations. Although ALTERNATING DECISION TREE and DECISION TREE performed relatively close on both data, their parameters differ substantially. For example, to classify the first part of negotiations ALTERNATING DECISION TREE builds 3 layers with 31 nodes, including 21 leaves, whereas DECISION TREE constructs 15 layers with 109 nodes, including 55 leaves. The same tree structures were obtained while classifying complete negotiations. SUPPORT VECTOR MACHINE' best performance is obtained when it uses linear polynomials and $C = 0.01$ to separate classes. RADIAL BASED FUNCTION NETWORKS first clusters the data into a small number of classes and then uses logistic regression to model them.

8 Importance of the Informativeness Features

Personal and possessive pronouns, e.g. *we*, *ours*, *you*, *yours*, are commonly present in the patterns with the informativeness words. We conduct a set of experiments in which 10 attributes corresponding to personal and possessive pronouns are added to the data representation. Table 8 presents the results. Again, *Acc*, *F*, *Pr*, *Rec* are computed by Eqs. 5–8; the results are estimated by tenfold cross-validation; $\beta = 1$. The baseline *Accuracy* equals 0.55 and *F* is equal to 0.71 which are less accurate than the reported machine learning results.

For most algorithms, adding pronoun attributes to the data representation either did not change or marginally changed the overall accuracy of classification. Only NEAREST NEIGHBOR substantially improved its classification results. ALTERNATING DECISION TREE and SUPPORT VECTOR MACHINE improved classification of successful negotiations while their correct classification of unsuccessful negotiations diminished. This holds for the first part of negotiations and complete negotiations. NEAREST NEIGHBOR significantly improved classification of successful negotiations without losing accuracy of classification of unsuccessful negotiations, although now it requires 25 neighbors to classify a data entry. DECISION TREE slightly decreased the number of nodes—to 95, including 48 leaves—when it classifies the first half of negotiations. However, information

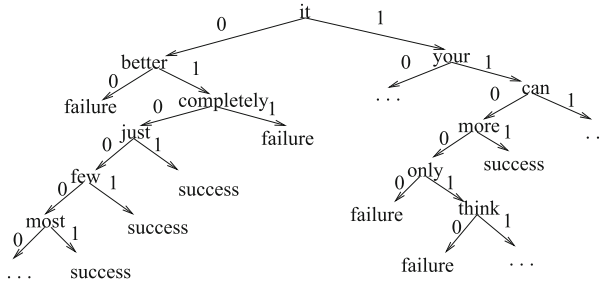


Fig. 1 Learning from texts of complete negotiations: the tree model of the language informativeness features for negotiation outcomes. More informative features appear higher on the tree. 1 shows the presence of the feature in a text, 0—its absence. For example, the path *it* = 1, *your* = 1, *can* = 0, *more* = 1 results in success. It means that when *it*, *your*, *more* are present, but *can* is absent in the record, there is a high probability of negotiation to succeed

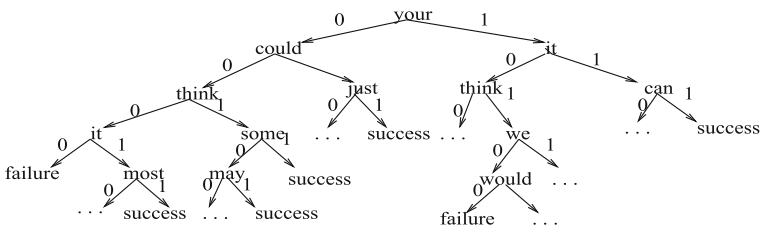


Fig. 2 Learning from texts of the first half of negotiations: the tree model of the language informativeness features for negotiation outcomes. More informative features appear higher on the tree. 1 shows the presence of the feature in a text, 0—its absence. For example, the path *your* = 1, *it* = 1, *can* = 1 results in success. It means that when *your*, *it*, *can* are present in the first half of the record, there is a high probability of complete negotiation to succeed

gain of individual features changes: when the most informative feature for complete negotiations is *it*, *your* is more informative in the first half of negotiations (Figs. 1 and 2). Another noticeable difference between comes from a heavier reliance on degree, scalar and comparative features. *Better*, *completely*, *few*, *just*, *more*, *most*, *only* are present at the top five layers of the model for complete negotiations, whereas in the model of the first half of negotiations only *just*, *most*, *some* are present.

We applied CONSISTENCY SUBSET EVALUATION (Hall and Holmes 2003) to find the smallest subset of features with a consistency level with class values equal to that of the full feature set.

$$Consistency_s = 1 - \frac{\sum_{i=0}^j |D_i| - |M_i|}{N} \tag{9}$$

where J is the number of distinct combinations of attribute values for s , $|D_i|$ is the number of occurrences of the i_{th} attribute value combination, $|M_i|$ is the size of the majority class for the i_{th} attribute value combination and N is the size of data set. We used Weka’s implementation of the algorithm, discussed by Witten and Frank (2005).

Table 9 Informativeness features: smallest subsets of features with the same consistency level with the class values as the full set of all features

First half	Complete
<i>we, I, you, our, my, your, they, it, its, just, only, extremely, over, quickly, maybe, quite, kind, bit, little, all, some, few, know, believe, think, must, may, can, should, would, could, same, than, most, more, further, better</i>	<i>we, I, you, our, my, your, they, it, almost, just, only, over, always, rather, maybe, perhaps, bit, little, all, some, few, know, believe, think, must, may, can, should, would, could, different, except, than, most, more, longer, further, better</i>

We consider such evaluation complementary because neither of the algorithms used it in their learning paradigms. The number of extracted words is approximately the same for two subsets—37 words for the first half of negotiations, 38 words for complete negotiations. It is interesting to note that the subset for the first half of negotiations contains more degree words *only, extremely, just, quickly, quite* than the subset of complete negotiations *just, only*. When both subsets contain comparatives *further, better*, the first subset contains comparative *same* and the second one—its opposite *different*. Table 9 reports the subsets.

9 Conclusions and Future Work

In this study, we have shown that *informativeness* of messages exchanged by negotiators correlates with negotiation success or failure. The analysis has been done for free-form, non-structured text records of electronic negotiations. We obtained the empirical evidence by conducting experiments on the Inspire text data.

Previously, studies of free-form text message informativeness for instance, by Carston (1998), have been conducted on a smaller scale, involving manual analysis of a restricted number of examples. In our work, Machine Learning methods allowed the analysis of a significantly larger number of examples. We analyzed language signals of informativeness provided by the presence or absence of degree, scalar, and comparative word categories. This representation was used in machine learning experiments to establish relations between informativeness and the negotiation outcomes. Using machine learning experiments on the first half of negotiations, we have shown that the informativeness signals may provide early prediction of the negotiation outcomes. We also showed that models of negotiation outcomes built from the informativeness features vary for complete and first half of negotiation records. All the reported empirical predictions exceed the baseline prediction of the negotiation success.

We see several directions for future work. Further analysis can be done to establish a time span which provides a reasonably reliable prediction of the negotiation outcomes. It would be interesting to establish relations between the span estimate and position within negotiation, e.g., the span may vary if it is estimated at the negotiation start and after introductions. In the future, we intend to analyze correlation between informativeness of messages and numerical values of the negotiation offers. This will help us to perform a *quantitative* analysis of relations between informativeness and

negotiation outcomes. Studying relations between factual and implied information revealed by negotiators and between explicit and implicit language used in negotiations are other promising venues for future research. The current work can be expanded by studies of emotions, e.g., by adding emoticon features to the text representation. Emoticons partially compensate the lack of visual information. Their connection with self-disclosure has been studied in McKenna et al. (2002). Outside language studies, the presented results can be used in training virtual agents for understanding traits of people behaviour. The agents can be trained on negotiation log files and use the obtained language features for prediction of negotiation outcomes.

Acknowledgments This work has been funded by the Natural Sciences and Engineering Research Council of Canada grants, available to both the authors. The authors thank Gregory Kersten for access to Inspire data and anonymous reviews for helpful comments.

References

- Adair W, Brett JM (2005) The negotiation dance: time, culture, and behavioral sequences in negotiation. *Org Sci* 16(1):33–51
- Carston R (1998) Informativeness, relevance and scalar implicature. In: Carston R, Uchida S (eds) *Relevance theory: applications and implications*. John Benjamins, Amsterdam, pp 179–236
- Chu-Carroll J, Carberry S (2000) Conflict resolution in collaborative planning dialogues. *Int J Hum Comput Stud* 53(6):969–1015
- Grice P (1989) *Studies in the way of words*. Harvard University Press, Cambridge, Massachusetts
- Hall M, Holmes G (2003) Benchmarking attribute selection techniques for discrete class data mining. *IEEE Trans Knowl Data Eng* 15(6):1437–1447
- Hine M, Murphy S, Weber M, Kersten G (2009) The role of emotion and language in dyadic e-negotiations. *Group Decis Negotiat* 18:193–211
- Kamakura W, Basuroy S, Boatwright P (2006) Is silence golden? An inquiry into the meaning of silence in professional product evaluations. *Quant Mark Econ* 4:119–141
- Kersten G, Zhang G (2003) Mining Inspire data for the determinants of successful internet negotiations. *Cent Eur J Oper Res* 11(3):297–316
- Koeszegi S, Pesendorfer E-M, Vetschera R (2007) Data-driven episodic phase analysis of e-negotiation. *Group Decis Negot* 2007 2:113–130
- Leech G, Svartvik J (2002) *A communicative grammar of English*. Longman, London
- Loken B (2006) Consumer psychology: categorization, inferences, affect, and persuasion. *Ann Rev Psychol* 57:453–485
- Mahmoudi MT, Badie K, Kharrat M (2008) Text organization via projection from researcher-space onto text-space. *Kybernetes* 37(8):1151–1164
- Manning C, Schutze H (2003) *Foundations of statistical natural language processing*. The MIT Press
- Marriot H (1995) Intercultural business negotiation, the discourse of negotiation. In: Firth A (ed), *The discourse of negotiation. Studies of languages in the workplace*. Pergamon, pp 247–268
- McKenna K, Green A, Gleason M (2002) Relationship formation on the internet: what's the big attraction?. *J Soc. Issues* 58(1):9–31
- Nastase V, Shirabad JS (2007) A study of sentiment and gender influence on negotiation outcome in electronic negotiations. In: *Proceedings of group decision and negotiations*, vol 2, pp 491–500
- Oakes M (2003) *Statistics for corpus linguistics*. Edinburgh University Press, Scotland
- Pennebaker J, Francis M, Booth R (2001) *Linguistic inquiry and word count: LIWC 2001*. Erlbaum Publishers, Mahwah
- Reitter D, Moore J (2007) Predicting success in dialogue. In: *Proceedings of the 45th annual meeting of the association of computational linguistics (ACL 2007)*. Association for Computational Linguistics, pp 808–815
- Roget's Interactive Thesaurus (2006) <http://thesaurus.reference.com/>
- Roloff M, Putnam L (1992) Introduction, communication and negotiation. In: Roloff M, Putnam L (eds) *Communication and negotiation*. Sage, London

- Sebastiani F (2002) Machine learning in automated text categorization. *ACM Comput Surv* 34(1):1–47
- Simons T (1993) Speech patterns and the concept of utility in cognitive maps: the case of integrative bargaining. *Acad Manag J* 38(1):139–156
- Sokolova M (2006) Learning from communication data: language in electronic business negotiations. Ph.D. dissertation
- Sokolova M, Szpakowicz S (2006) Language patterns in the learning of strategies from negotiation texts. In: Proceedings of the 19th Canadian conference on artificial intelligence (AI'2006), pp 288–299, Springer
- Sokolova M, Lapalme G (2007) Informativeness for prediction of negotiation outcomes. In: Proceedings of group decision and negotiations 2007, vol 2, pp 501–511
- Sokolova M, Nastase V, Szpakowicz S (2008) The telling tail: signals of success in electronic negotiation texts. In: Proceedings of the third international joint conference on natural language processing (IJCNLP 2008) pp 257–264
- Sperber S, Wilson D (2006) Pragmatics. In: Jackson F, Smith M (eds) *Oxford handbook of contemporary philosophy*. Oxford University Press, Oxford
- Takeuchi H, Subramaniam L, Nasukawa T, Roy S (2007) Automatic identification of important segments and expressions for mining of business-oriented conversations at contact centers. In: Proceedings of the 2007 joint conference on empirical methods in natural language processing and computational natural language learning (EMNLP-CoNLL) pp 458–467
- Witten I, Frank E (2005) *Data mining: practical machine learning tools and techniques*. Morgan Kaufmann, Los Altos, CA
- Yu G (2010) Lexical diversity in writing and speaking task performances. *Appl Linguist* 31(2):236–259