Human Language Technology for Text-based Analysis of Psychotherapy Sessions in the Spanish Language

Horacio Saggion[α], **Elena Stein-Sparvieri**[β], **David Maldavsky**[β], **Sandra Szasz**[γ]
[α]DTIC - Universitat Pompeu Fabra
Calle Tanger 122-140, Poble Nou
Barcelona - Spain
H.Saggion@dcs.shef.ac.uk
[β]IAEPCS - Universidad de Ciencias Sociales y Empresariales
Paraguay 1401, PB, Bs. As. Argentina
estein@solutiion.com.ar;dmaldavsky@elsitio.net
[γ]University of Sheffield - Department of Computer Sciences
211 Portobello Street - Sheffield - UK
S.Szasz@sheffield.ac.uk

Abstract

We present work in progress in the application of Natural Language Processing (NLP) technology to the analysis of textual transcriptions of psychotherapy sessions in the Spanish Language. We are developing a set of NLP tools as well as adapting an existing dictionary for the analysis of interviews framed on a psychoanalytic theory. We investigate the application of NLP techniques, including dictionary-based interpretation, and speech act identification and classification for the (semi) automatic identification in text of a set of psychoanalytical variables. The objective of the work is to provide a set of tools and resources to assist therapist during discourse analysis.

1 Introduction

Computer-based textual analysis in psychology is not new; in psychotherapy, electronic dictionaries and other lexical resources are widely used to analyse both therapist's and patient's discourses produced during psychotherapy sessions. In this paper we present work in progress in the application of Natural Language Processing (NLP) technology to the analysis of psychotherapy sessions in the Spanish Language. Based on a psychoanalytic theory, we are developing a set of NLP tools as well as adapting an existing dictionary for the analysis of interviews. We investigate the application of NLP techniques, including dictionary-based interpretation, and speech act identification and classification for the automatic analysis of spoken transcriptions in Spanish of psychoanalysis sessions between therapists and patients. In Figure 1 we show a fragment of a manually transcribed interview in Spanish (and its translation to English) from our development corpus.

The automatic analysis of the sessions, which is used as a tool for assessment and interpretation of the transcribed psychotherapy sessions is based on a theory developed by Liberman and extended by Maldavsky (Liberman and Maldavsky, 1975) and framed on Freudian theory (Freud, 1925). The automatic tools to be presented here aim at recognizing a subset of *Freudian drives* manifested in both patient's and therapist's discourse.

The objective of the analysis is not to provide a full automated solution to discourse interpretation in this area, but a set of tools and resources to assist therapists during discourse analysis. Although work in text-based interpretation in psychology is not new, researchers in our project have identified limitations in current practices due to the fact that current text-based systems do not tackle ambiguity problems at lexical, syntactic, or semantic levels: for example systems that consider out-of-context superficial forms would are unable to distinguish between different used of the same lexical item ("para" as a preposition vs. "para" as a form of the verb "parar" (to stop); "rio" as a common noun vs. "rio" as a contextual clue for the identification of a geographical name; etc.). The use of advanced natural language processing techniques could help produce

Proceedings of the NAACL HLT 2010 Young Investigators Workshop on Computational Approaches to Languages of the Americas, pages 32–39, Los Angeles, California, June 2010.

Transcribed Session (Spanish/English Version)
T: ¿con que te cortaste? (T: What did you cut yourself with?)
L: con un vidrio que encontré en el patio (L: With a glass I found in the patio.)
T: ¿donde lo tenías? (T: Where did you have it?)
L: en el locker, en la puertita del locker, y después lo puse en la jabonera cuando baje a bañarme (L: In the locker, in the locker's small door, and then I put it in the soap box when I went down to have a bath.)
T: o sea, ya tenías un vidrio escondido (T: so, you already had the glass hidden.)
L: síı, ayer lo encontré (L: Yes, I found it yesterday.)
T: ¿ayer a la tarde? (T: Yesterday afternoon?)
L: sí, sí, de ayer a la tarde (L: Yes, yes, yesterday afternoon.)

Figure 1: Transcription of a small fragment of a therapy session in Spanish and its translation to English. **T** indicates therapist and **L** indicates patient.

better analysis of the input material and therefore be used for a better diagnosis and follow-up. It is worth mentioning that full interpretation of therapy sessions is not only based on textual analysis, but also in other elements of the session such as the actual speech (e.g. pitch), para-verbal elements such as patient movement, etc. This work addresses only text interpretation issues.

The rest of the paper is organized as follows: Section 2 describes related work in the area of computational tools for text analysis in psychology. In Section 3, the theoretical framework for our work is briefly introduced. Section 4 describes the implementation of NLP tools for the analysis of the interviews and Section 5 closes the paper describing current and future work.

2 Related Work

There are a number of well-established computational tools for the analysis and extraction of meaning from text in the social sciences (See (Alexa and Zuell, 2000) for an overview of tools and resources). Some tools are bound to particular theoretical principles, for example the LWIC dictionary (Pennebaker et al., 2001) encodes specific categories to be identified in text while others follow a theory-free approach (Iker and Klein, 1974) where the theory emerges from the analysis of the data.

There has been substantial research in the development of methods to analyze linguistic input in the field of psychotherapy in order to measure a number of psychological variables such as emotion, abstraction, referential activity, etc. among them Bucci's Referential Activity (RA) non-weighted (Bucci, 2002) and weighted dictionaries (Bucci and Maskit, 2006) for the English language, or Höltzer and others' affective dictionary (Hölzer et al., 1997) for the German language. The LIWC tool has been used to detect different types of personalities in written self-descriptions (Chung and Pennebaker, 2008). This program counts meaningful words that express emotion, abstraction, verbal behavior, demographic variables, traditional personality measures, formal and informal settings, deception and honesty, emotional upheavals, social interaction, use of cognitive and emotion words, word analysis in psychotherapy, references to self and others. For Spanish (Roussos and O'Connell, 2005) have developed a dictionary in the area of psychotherapy

to measure referential activity.

Early work on dictionaries in the area of psychology include the General Inquirer psychosociological dictionary (Stone and Hunt, 1963) which can be used in various applications; current work on lexical resources for identifying particular text variables – such as measuring strong/weak opinions, sentiments, subjective/objective language, etc. – include the SentiWordnet resource (Esuli and Sebastiani, 2006) derived from WordNet which has been used in various opinion mining works (Devitt and Ahmad, 2007); other lines of research include the derivation of word-lists (semi) automatically for opinion classification (Turney, 2002). To the best of our knowledge, little research has been carried out on natural language processing for discourse interpretation in psychology.

3 Theoretical Framework Overview

Liberman's theory identifies 7 drives (i.e., a subset of Freud's drives) which are introduced in Table 1 we may associate these drives with emotional or affective states such as: strong emotions associated with IL; ecstasy or trance with O1; sadness with O2; anger with A1; concrete language with A2; warnings, suspense, and premonition with UPH ; and congratulation, adulation, and promises with GPH. In diagnosis these variables are associated to pathologies such as addiction, schizophrenia, depression, paranoia, obsession, phobia, and hysteria; so their manifestation in text is of paramount importance for diagnosis.

Abbreviation	Drive Name
IL	Intra-somatic libido
O1	Primary oral
O2	Secondary oral sadistic
A1	Primary anal sadistic
A2	Secondary anal sadistic
UPH	Urethrae phallic
GPH	Genital phallic

Table 1: Drives in Liberman and Maldavsky theory

The theory also associates lexicalizations to each of the drives (Maldavsky, 2003), thus creating a semantic dictionary with 7 categories, the main working hypothesis is that drives manifest through linguistic style, present at word level, phrase, and narrative. Lexicalisations for each drive have been carefully selected following a variety of methods including manual derivation of words from concepts, study of texts where a scene is clearly present (e.g., everyday activities), use of thesaurus, etc. Ambiguity is preserved and a lexicalisation can signal more than one drive. We show some lexicalisations in Table 2.

Drive	Lexicalisation
IL	**verbs**: to throw up, to break; **nouns**: hospital, throat; **adjectives**: sick, fat; **adverbs**: fatally, greedily
O1	**verbs**: to sip, to suck; **nouns**: enigma, research; **adjectives**: mystical, enlightening; **adverbs**: elliptically, enigmatically
O2	**verbs**: to feel, to feel like; **nouns**: feeling, victim; **adjectives**: sensitive, happy, sad; **adverbs**: fondly, obediently
A1	**verbs**: to bother, to kick; **nouns**: violence, transgression; **adjectives**: angry, locked; **adverbs**: angrily,boldly, crossly
A2	**verbs**: must, to know; **nouns**: vice, doubt; **adjectives**: good, bad; **adverbs**: but, although, however
UPH	**verbs**: to be able, to dare; **nouns**: scar, precipice, wound; **adjectives**: coward, scared; **adverbs**: almost, a bit
GPH	**verbs**: to promise, to give; **nouns**: beauty, ugliness; **adjectives**: wavy, pretty; **adverbs**: more, even

Table 2: Sample of drives and associated lexicalisation

In addition to word-level analysis, the theory provides methods for analysis at narrative and speech act level.

Speech acts are actions performed when making an utterance (Searle, 1969) and they include (Searle, 1976) illocutionary (e.g. assert, suggest), perlocutionary (e.g. convince, insult), and propositional (e.g. making a reference) types. There has been substantial work on speech act segmentation and classification. Different authors adopt different classifications or theories of speech acts in order to restrict the categories to those relevant for the purpose of analysis. For example, in dialogue systems (Allen et

Drive	Speech Acts
IL	references to the state of things; reference to body and body processes; etc.
O1	abstract deduction; negation; reference to physical disomfort; etc.
O2	lamentation; complain; beg; etc.
A1	verbally abuse; provoke; confront; etc.
A2	judge; clarify; confirm; etc.
UPH	forewarning; warning; inquest; counsel; etc.
GPH	congratulate; thank; promise; exaggerate; etc.

Table 3: Drives and Speech Acts

al., 1996; Henry Prakken, 2000), the list of speech acts may vary from 4 to 10 categories and it may include acts such as assertion, WH-question, directives, greeting, direct/indicrect request, etc.
The pychoanalytic framework we are following has its own inventory of speech acts. The objective is also to link scenes in narratives and speech acts to the 7 drives (in Table 1). There is a variety of speech acts in the adopted framework, in Table 3 we present a sample of speech acts associated to each of the drives. The objective of the semi-automatic analysis is to help their identification to facilitate the work of the psychotherapist.

4 Text Analysis of Interviews

We have implemented a series of programs, lexical resources, and grammars to process interviews and other types of textual data in Spanish. We are using the GATE system (Maynard et al., 2002) as an infrastructure or development framework; most developments are new, not included in the GATE system, and they are packaged in a plug-in which can be accessed through the GATE system or used standalone. We have developed various programs to automatically annotate the interviews including segmentation of the transcription, word-based thematic segmentation, tagging, and dictionary-based interpretation and analysis.

4.1 Dictionary

One of the main components of the system is a dictionary which is taken as the basis for text interpretation. This is being implemented as a language resource in GATE. It is based on lists of word forms which have been created for each of the drives. The lists are organized according to their parts of speech. The available dictionary (Maldavsky, 2003) contains all inflected forms of verbs, nouns, adjectives, and adverbs which we are transforming into a dictionary which will contain only roots. An instance of the dictionary is created from the set of lists and kept on-line for processing. The current version of the dictionary (inflected forms) contains over 298 thousand verb forms, over 22 thousand noun forms, over 137 thousand adjectives, and over 9 thousand adverbs. An annotation tool has been implemented based on a schema for our dictionary, we use the graphical user interface functionalities provided by the GATE infrastructure allowing a researcher annotate words she may want to included in the dictionary or segment the text in units for further analysis.

4.2 Programs for Interviews' Interpretation

The following programs used for the automatic analysis of the interviews.

- A wrapper to the TreeTagger parts of speech package (Schmid, 1995) (See `http://www.clarin.eu/tools/treetagger`) has been implemented in order to call it from the GATE system and an alignment program has been developed to associate the output of the tagger to the actual text of the interview, therefore creating word annotations containing features from the TreeTagger and additional features computed by our programs. Note that the TreeTagger distributed with GATE was innapropriate for our purposes because it does require tokenisation of the input performed before invoking the tagger, this is the reason why we had to create our own wrapper.

- A segmentation program is used to identify patient and therapist interventions.

- Text chuncking and named entity recognition is being developed using Support Vector Machines and training data from the CoNLL

evaluation program. We have created a trainable system using machine learning resources provided by the GATE framework. The CoNLL 2002 Spanish dataset which provides information on named entities such as *Location*, *Organization*, *Person*, and *Miscellaneous* was analyzed using parts-of-speech tagging, morphological analysis, and gazetteer lookup in order to derive a set of features for learning. A supports vector machine was trained that uses gazetteer information, word level information, orthography, parts-of-speech, and lemmatization. We have collected a number of lists to assist the identification of names of organization, persons, locations, time expressions, etc. The performance of the current system is at 68% F-score. Note that named entity recognition is particularly important to track names in longitudinal analysis of interviews, but also to disambiguate names which in Spanish are ambiguous (e.g. "amado" can be a person name in addition to a form of the verb "amar"; "quito" can be the name of a place in addition to a form of the verb "quitar", etc.)

- A program uses the dictionary and interprets each word or complex term according to the drives in the dictionary taking into account parts of speech information and named entity recognition.

- A topic segmentation program has been implemented to break the interview in fragments which can be selected for fine-grained interpretation. This module is based on tf*idf similarity between candidate segments. A second module we are implementing aims at the recognition of segments referring to prototypical scenes a patient may refer to: family, work, love, health, money, etc. Further gazetteer list information has been collected from Spanish sources to create lexicons for assisting the automatic identification of the above categories. We are in the process of manually annotating a set of transcriptions as the basis for training a classification system for this task. Conceptual information will be used for this purpose.

- A processing resource has been implemented to generate an interpretation of the different languages or drives' variables for different segments chosen by the human analyst (therapist or patient or any other segment of interest) and statistics are computed for each of the segments; these can be exported for the therapist to carry out additional analysis and interpretation. Note that the current tool considerably improves the previous practises in dictionary-based interpretation, since the implemented tool takes into account syntactic and semantic information as a filter for interpretation.

4.3 Rule-based Speech Acts' Detection

We are carrying out induction sessions with psychotherapits in order to capture ways in which speech acts in the adopted framework are expressed. The induction sessions provided valuable material to start implementation of a rule-based speech act detection program (with regular expressions and a dictionary) based on use use of syntactic and lexical information. These procedures allow us to collect a set of expressions and lexical/syntactic patterns for objective identification of a subset of speech acts. We are also annotating the development corpus of interviews (a total of 30 will be annotated with a minimum of 2 annotators per interview) with speech acts categories. Each speech segment is annotated with one main speech act and a number (possibly zero) subordinate speech acts. We are using the GATE environment to provide appropriate support for the annotation process. In Figure 2 we show a fragment of interview in the annotation tool annotated according to the interpretation of one of our judges (the annotation window shows a "complaint" speech act associated to the fragment "no me estaba tratando de entender como él siempre hace" ("he did not understand as he always does")). We expect the annotated corpus to be a valuable resource for the development of a trainable speech act recognition program based on lexical clues and syntactic infor-

mation. This trainable system will extend the rule-based approach or incorporate the rule-based analysis into it.

A sample of expressions we have identified and implemented for a subset of speech acts is presented in Table 4. The analysis of speech acts will provide an additional level for drive's identification.

5 Perspectives and Current Work

We have described our initial work on a set of tools being developed for the analysis of psychotherapy interviews in the Spanish language. The tools extend work on dictionary-based text interpretation by incorporating NLP tools such as tagging, topic/scene segmentation, speech act detection, and named entity recognition. One main contribution of our research is the implementation of a dictionary for the Spanish language which can be used not only for the identification of Freudian variables but also for work on affective language and sentiment analysis. We are currently working on the development of a full module for speech-act recognition and on the creation of a corpus of annotated interviews which will serve for further training and evaluation purposes. The set of resources developed in the project will be made available to the computational linguistics community for research purposes. We think that although this is work in progress it is worth mentioning evaluation. Where evaluation of the tools is concerned, we are carrying out intrinsic evaluation comparing annotated categories against predicted categories currently for named entity recognition and discourse segmentation and in the future for speech act recognition and classification. Where more extrinsic evaluation is concerned, we will evaluate how the tools presented here can help theraphist in better interpretation of clinical data. The implemented tools will also be used to compare word-level based interpretation produced by the dictionary to interpretation produced by the analysis at speech act level.

Acknowledgements

We thank the reviewers for their very useful comments. This work was partially supported by a grant from the Royal Society (JP090069), UK. The first author is grateful to Programa Ramón y Cajal 2009 from the Ministerio de Ciencia e Innovación, Spain.

References

M. Alexa and C. Zuell. 2000. Text analysis software: Commonalities, differences and limitations: The results of a review. *Quality & Quantity*, 34:299–231.

J. F. Allen, B. W. Miller, E. K. Ringger, and T. Sikorski. 1996. A robust system for natural spoken dialogue. In *Proceedings of the 34th annual meeting on Association for Computational Linguistics*, pages 62–70, Morristown, NJ, USA. Association for Computational Linguistics.

W. Bucci and B. Maskit. 2006. A Weighted Referential Activity Dictionary. In *Computing Attitude and Affect in Text: Theory and Applications*, volume 20 of *The Information Retrieval Series*, pages 49–60. Springer Verlag.

W. Bucci. 2002. Referential Activity (RA): Scales and computer procedures. In *An Open Door Review of Outcome Studies in Psychoanalysis*. International Psychoanalytical Association.

C.K. Chung and J.W. Pennebaker. 2008. Revealing dimensions of thinking in open-ended self-descriptions: An automated meaning extraction method for natural language. *Journal of Research in Personality*, 42:96–132.

Ann Devitt and Khurshid Ahmad. 2007. Sentiment polarity identification in financial news: A cohesion-based approach. In *Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics*, pages 984–991, Prague, Czech Republic, June. Association for Computational Linguistics.

Andrea Esuli and Fabrizio Sebastiani. 2006. SENTIWORDNET: A publicly available lexical resource for opinion mining. In *Proceedings of LREC-06, 5th Conference on Language Resources and Evaluation*, pages 417–422, Genova, IT.

S. Freud. 1925. *Obras Completas*. Amorrortu (Eds.), Madrid, Spain.

H. Henry Prakken. 2000. On dialogue systems with speech acts, arguments, and counterarguments. In *Logics in Artificial Intelligence*, pages 224–238. Springer Verlag.

M. Hölzer, D. Pokorny, H. Kächele, and L. Luborsky. 1997. The Verbalization of Emotions in the Therapeutic Dialogue-A Correlate of Therapeutic Outcome? *Psychotherapy Research*, 7(3):261–273.

H.P. Iker and R. Klein. 1974. WORDS: A computer system for the analysis of content. *Behavior Research Methods and Instrumentation*, 6:430–438.

D. Liberman and D. Maldavsky. 1975. *Psicoanlisis y semitica*. Paidos, Buenos Aires, Argentina.

D. Maldavsky. 2003. *La investigacin psicoanaltica del lenguaje: algoritmo David Liberman*. Editorial Lugar, Buenos Aires, Argentina.

Speech Act	**Pattern or Expression**
beg	PPX + *rogar* | *implorar* | *suplicar*
demand	PPX + *exhortar* | *exigir* | *demandar* | *perdir*
demand recognition	*decir* que esta bien | correcto | perfecto | bueno; está bien, no?
demand forgiveness	PPX + *perdonar*
justify	por que; por eso; debido a que; por esa razón
permission	con *PPO* permiso; *pedir*; PPX + *dejar*
interrupt	para... para; espera...; ah me olvide...
cite	como dijo NP | PPX ; según NP | PPX ; de acuerdo con NP | PPX
sinthesis	en resumen; para concluir; en síntesis
doubt	no PPX *quedar* | *ser* | *estar* claro; quien sabe
trust/distrust	no *confiar* | *desconfiar*; *confiar* | *desconfiar*
submission	*tener* razón; no + PPX + *enojar*
appeal	decime que me querés; ...
compassion/self-compassion	me da pena; pobre; pobrecito;...
sacrifice	yo que hice todo esto; yo que te di todo; si no fuera por mi; ...

Table 4: Speech Acts and Lexical/Syntactic Patterns (PPX = pronouns; NP = proper nouns; PPO = possessive)

D. Maynard, V. Tablan, H. Cunningham, C. Ursu, H. Saggion, K. Bontcheva, and Y. Wilks. 2002. Architectural Elements of Language Engineering Robustness. *Journal of Natural Language Engineering – Special Issue on Robust Methods in Analysis of Natural Language Data*, 8(2/3):257–274.

J.W. Pennebaker, M.E. Francis, and R.j. Both. 2001. *Linguistic Inquiry and Word Count (LIWC)*. Erlbaum Publishers.

A. Roussos and M. O'Connell. 2005. Construcción de un diccionario ponderado en español para medir la Actividad Referencial. *Revista del Instituto de Investigaciones de la Facultad de Psicología - UBA*, 10(2):99–119.

H. Schmid. 1995. Improvements in part-of-speech tagging with an application to german. In *In Proceedings of the ACL SIGDAT-Workshop*, pages 47–50.

J. Searle. 1969. *Speech acts: An essay in the philosophy of language.* Cambridge University Press.

John R. Searle. 1976. A classification of illocutionary acts. *Language in Society*, 5(1):1–23.

P. J. Stone and E. B. Hunt. 1963. A Computer Approach to Content Analysis: Studies using the General Inquirer System. In *Proceedings of the Spring Joint Computer Conference*, pages 241–256, New York, NY, USA. ACM.

P. D. Turney. 2002. Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews. In *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics (ACL '02)*, pages 417–424, Morristown, NJ, USA, July. Association for Computational Linguistics.

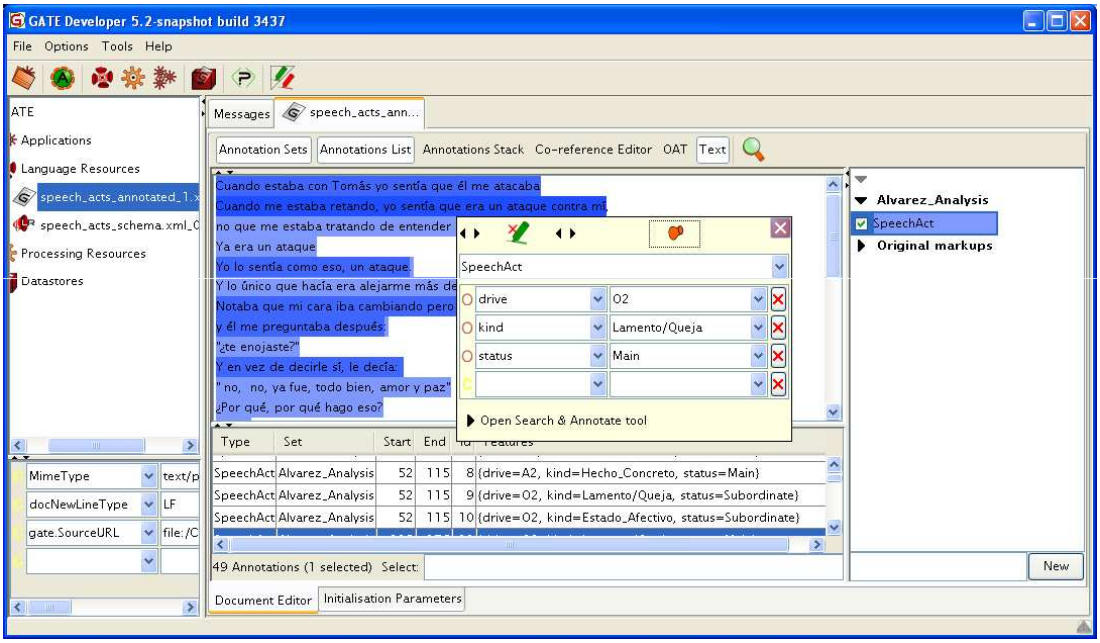

Figure 2: Speech Acts Segmentation and Interpretation