1. Introduction

Studying discourses seems to be an endless occupation, mostly because of the different approaches to the concept of ‘discourse’, but also because of the ever changing media and contexts where discourses take place. Our attention has always been drawn to the idea that discourses are social interactions which leave traces in our lives. Thus, we feel close to the critical discourse analysis (CDA) approaches on the analysis of such interactions, mainly to the detection of the social profile of the actants, and the tools that seem to be the most appropriate to our goal are, in our opinion, the pragmalinguistic markers present in discourses. As a new kind of medium – the Internet – is gaining over the others (or it only seems to gain!), we will therefore pay attention to what happens there, too. Being a study trying to determine the level of chances one could have to establish patterns in such a dynamic activity like the discourse, we don’t think of identities, but of generic profiles, types of individuals, as they come on the basis of their habits in linguistic communication. As it is not (yet!) an applied study, there will be no corpus exploited in it, and the method will largely resemble a contemplation of common sense observations already made in discourse analysis of all kinds and in the field of pragmalinguistics. We are just at the beginning of an attempt to verify whether such (discursive) social profiles as those described above can be extracted by means of automatic tools (NLP tools) from discursive interactions taking place online. Journey with us and be prepared to share our joy as to whether the conclusions will be satisfactory!

They say the utterer is always right, because s/he knows why s/he has made the choices that s/he has made from the infinite offer of the language to express her/his thoughts or intentions. Anyone of us humans is entitled to believe they are right, because they know better than anyone else what they meant to say. We could then consider that one can always rely on her/his own sayings, on the linguistic expression of their respective thoughts or intentions. People can unveil or dissimulate whatever they want in their verbal interventions. They have control over their own expressions. Or haven’t they? Nevertheless, pragmalinguistics and the CDA show that there are tools to detect more than it is said in a given verbal expression (oral or written). Markers in discourse are giving out lots of information no-one would ever agree to unveil of themselves for different reasons. From all the elements that describe linguistic competence, pragmalinguistics deals with the concept of "choice". The choices we make are describing us as social actors in different contexts. Observing this could offer ideas for new approaches in the study of language use, with or without automatic tools to help us in our efforts.
2. The reality of language use and some theoretical foundations

Let us start from a reality of all languages, because it is a language universal: there are no such things like identical words (semantically speaking) in one and the same historical language, there is no perfect synonymy. This brings us directly to the issue under discussion: trying to express ourselves, we make choices. The vocal of Simply Red does not say "I hope you'll understand", but "I know you'll comprehend", because it is not about reasoning, coping with some already existing bit of information, but capturing and processing new information in order to make it part of the existing knowledge. Rational vs. cognitive processing. Apparently, any average native speaker of English would assume it’s the same thing. And it is. Up to a point! The above distinction between the semantics of the two words can stand for an example for what I am trying to talk about in the present study. People would say: potato-potato, but some specialist could find not necessarily the semantic distinction as such, but what it could make us think about. Some of us would say "comprehend" is a mark of presumptuousness, while "understand" is the "normal" way of speaking. Some would speak of the need of belonging (to some elite who use fancy words to say even trivial things). There will also be people for whom the meaning of "comprehend" is unknown, so they will simply not understand a thing. Looking to the facts everybody could remark in the everyday life, we can add situations such as the one where we have an individual using a "fashionable" word without knowing its exact meaning. I knew someone who used to say (frequently!) "mass media" instead of "average people", and we all know at least one person who says "fortuitous" (casual, accidental) with the meaning of "forced", just trying to use a fancier word and by that impress her/his audience.

This is not all. Let us now think of faking a mistake, for fun or for educational purposes. And then let’s think of those kinds of situations where the utterer knows what s/he is saying, but in the end, s/he realizes s/he’s made the wrong choice of words with respect to the interlocutor. Here we have touched a fundamental idea related to the discourse: it takes (at least) two to have one! There is more: each of the two interlocutors has to try their best to get a representation of the other as near as possible to the reality of the other.

Getting back to the core issue of this study, we have to accept the idea that even adequacy is a question of choice (which could be good or not so good). We shall look now to some possible conclusions an analyst could extract from discursive production of known but more than this of unknown people. We are talking of some kind of second degree inference, and we have to bear in our minds that in doing so we use (almost) everything, from the physical context of the production of the discourse, to the common history of the two and the contract of communication they have agreed upon for that particular instance of discursive activity. The scariest part in all this is what Ivan Preston captured in his phrase saying that “meanings are in people, not in messages”\(^1\).

Almost everything can be recomposed – a set, a town, the world itself, even the climate! – but there is no such thing like the recreation of a state of human nature. Not only are we different from one another, but no individual can ever repeat his same self. This spoils the context, and makes it more than difficult to have rigorous patterns of human behaviour. And yet! Psychologists talk about the fact that where there is human activity, there are patterns. Obviously, they settle for less, but this is the only chance to make some progress in interpreting human behaviour. This is the chance also for all IT approaches on the study of language uses. As approximate as they can be, the patterns reproduce that part of the language uses which remains somehow the same in different situations (which have not to be so different, but seemingly similar!). Looking back to the examples above, we can see that there can be more than one interpretation of the use of some word ("comprehend") instead of its (helas!, imperfect) synonym "understand". Going further, the question would be whether we can "capture" the personality of the utterer of a phrase like the one in the song by Simply Red, by the simple fact that s/he made a specific choice of words. Could that option be a sign of ignorance? Or a sign of a good knowledge of the English language? Or maybe a desire to make some impact on the listener? Or…? Yes, yes and yes! All of the above. And more.

3. **Important authors to keep in mind**

In one of his published books (Marcus, 2006), Solomon Marcus talks about imprecision in language use. I dare anyone to find a better term to name the phenomenon! One could find in that book discussions on many kinds of imprecision, such as: the paradox, the antinomy, the independence, the undecidability, the incompleteness, the axioms of choice, the fractals, the plausibility (with its special case, the possibility) and indeed many others. The presence of an imprecision is difficult enough to manage in trying to get the meaning of some discourse, but Solomon Marcus invites the reader to think of even more complicated cases of interference of different imprecision: as significant examples, we have the way random interferes with negligible, the way chaotic interferes with negligible, or the random with the gradual. It is quite alarming. We could ask ourselves how do we still manage to communicate using language, when we cannot really rely on what we are saying or on what we think we understand from what others tell us. It seems that using language is one of the riskiest activities, but we continue to pretend to understand each other in this way. There is more: we like to communicate with the help of this marvelous tool, and we love to play with it, to ex-press ourselves, or to dissimulate ourselves using it. Using language is pretending. It’s sending around hints to direct our fellows humans in their effort of understanding what we are expecting them to do. It works. Not always, not perfectly, not with everybody, not in every circumstances. The imprecision make the beauty of the language, but they also make us look like victims of our own "habit" to speak.

The aim of the present study is to look a little longer at the later aspect: us as victims, in this particular sense of the language unveiling to the analysts things we wouldn’t want to go public. The main target at the present time is the discourse on the Internet and the perspective is that of the analyst. The challenge is bigger, because the method should be new, the environment is new (and still not quite known or understood), the people interfering there are numerous (not actually, but as declared identities), the manner of using language is at too many levels of (im)perfection, and above all the freedom is unlimited. To keep this tremendous freedom and to really enjoy it, people tend to hide
their identity. Looking closer at the phenomenon we could remark that this is not such a new behaviour. As we play roles all the time in society, we pretend. It’s true that in the actual world, where we can bump into one another if we want to, getting to know who you are interacting with is easier and more accurate than in an environment which is based on make believe. To give a name to this problem, we could say that it is an issue of representation and metarepresentation, which is the psycho-social foundation of human communication of every kind. The better you can represent your interlocutor in your head, the faster you will be able to tune your verbal intervention to match his/her expectations, and consequently your discourse will be more effective. We know it’s difficult enough to do that in the actual world, and that we often fail. So, what are the chances to do it right on the Internet? Practically zero! Should we just give up? No! We will just change our focus: we won’t go after individuals, but after types of individuals, and this is possible because individuals – in actual life or in cyberspace – copy each other, they follow fashions, they imitate. So, they have become to resemble one another, and one could create groups of profiles describing common behaviour of various groups of people. It is a way of creating typologies in a world that refuses to give us enough information, but which is too important for either of us to ignore. All the analyst should have to do is to pick up elements of discourse that could speak of the speaker. The fundamental tool will come from pragmalinguistics (Veltman, Steiner, 1988), and it will be, as shown above, the choice any speaker makes in constructing their discourses. It resembles to the logical syllogism: you choose that, you are like this (All those who choose X are of the class Y; the individual speaker N chose X; then N is of the class Y). Classes (or types, or generic profiles) are to be predetermined, in a focused research, starting from the interest we focus on in the study. If the interest is just to "see" who is there, on the Net, classes could form along the way. Let’s think, for example that it was never hard to discriminate between those with something like a real name (say, Malcolm Gladwell) and those calling themselves "angry bird", or "big boy", or something of the sort. We shall never really know that "Malcolm Gladwell" is the actual name of the one who is using it on the Internet, but isn’t this a problem even in the concrete world we live in?! So therefore we will be more inclined to believe that an individual who used a "real" name is in fact the person having that particular name. Can we now form a class with "real names" and another with nicknames? Yes, we can! This was just an example, but it is enough to prove the point. So, starting from the declaration of their respective identities, the forums’ users can be broken down in classes like: complete real name, partial real name and nickname (here, we could have many subclasses as the nickname varies from highbrow to low brown). One could also imagine more classes and subclasses, but the point is that we can distinguish patterns in the behaviour of the writers on the Internet from this point of view. What we have to avoid will be the temptation to transform patterns into rules!

Then, we can have as a criterion the capacity to stay with the idea of the article, to keep close to the topic the journalist has proposed in his column. A study of isotopies (surface ones as well as deep ones) would show how close to the article topic is a comment or a comment of a comment. Based upon this, classes of profiles could be drawn: on topic, close to the topic, vaguely close to the topic (with many subtypes: those who do not understand what the article is about, those who understand, but who would want to change the topic by keeping the same context with the same interlocutors, those who understand a part of the topic, but who are more interested in the author), not even close to understanding, wishing just to be present online, but also
some kinds of *ad persona*, meaning they do not argue on the topic but attack the journalist or the writer of a previous comment. Then, we have the educated ones vs. the not so educated ones, the aligned ones vs. the neutral ones, the rational ones vs. the emotive ones and many other couples of opposite profiles.

As we can see, a lot of information can be extracted from the actual expression a writer gives to his/her thoughts or intentions. We are talking of information on the writer, something giving him away. Jean-Pierre van Noppen is talking of *tenors*, which are to be taken as characters or patterns of something (van Noppen, 2009). Such a tenor can tell what kind of person the author is or is pretending to be, what kind of people the expected audience are, what the relationship between them is (or what relationship the authors is presenting it as). Tenors could be: relative status (equality or inequality between interlocutors), social distance (eg. familiarity, friendliness), personalization, standing (or how much the author comes across as an authority on the subject), stance (broken down further into attitude and modality, which can be epistemic or deontic).

Starting from those tenors, we can analyze and construe lexical and grammatical choices of a given author and conclude to some pattern s/he could belong to. Following the description of that pattern, a class of individuals can be considered as opposed to other classes with different descriptions. I will mention an example from the book by Jean-Pierre van Noppen, to illustrate this idea: Explaining the tenor called *stance*, the author presents several kinds of attitudes, more or less explicit, and therefore more or less easy to detect. Let’s have a look at the paragraph:

" Attitudes […]:

- Asserted attitudes are attitudes which are mentioned quite openly, a typical reader is aware of them and is free to disagree with them: "The government’s behaviour was disgraceful".

- Assumed attitudes are attitudes which are mentioned as if they were truths accepted by everyone, on which an argument can be built: "After nine years of the government betrayal, …" (Main argument follows). A typical reader will feel less free to disagree with assumed attitudes.

- Triggered attitudes aren’t mentioned at all, but a typical reader will infer them. Example: "Even though Fred’s father is very old, Fred only visits him once a year”. This triggers a negative attitude towards Fred. […] the syntax employed ("even though … only") encourages a typical reader to infer from the facts a negative attitude towards Fred.

One way of triggering attitudes is the manipulation of agency and affectedness in a text: wording material processes in such a way that certain entities appear as actors (and therefore come across as responsible for what happened) while others appear as goals (and therefore come across as more or less victims of what happened). " (van Noppen, 2009).
4. More authors and practical discussions

When it comes to expressing extreme nationalism or xenophobia, an author can put it bluntly in the text or it can simply make some unnecessary precisions like the French journalist who reported a crime committed somewhere in Paris mentioning that the perpetrator was a young French citizen of Moroccan extraction. If he would have simply said that it was a young man living in Paris, the missing accent on the ethnicity of the criminal wouldn’t have created the possibility to infer that immigrants are a problem for the native French people. It goes similarly for precisions like "a Roma Romanian was caught while she...". An analyst can easily understand that the crime itself is not the major preoccupation of the journalist, but the immigrants are. The same goes for the comments’ authors or for the authors of comments to other comments. We could imagine worse: what if the journalist (or some commentator on the forum) says the same thing in another form (like: there is no point to ask ourselves who the perpetrator is, because it has become common knowledge that Maghrebians/Roma Romanians do these kind of things to us)? What are we going to feel is the obvious conclusion? Would we be able to remember the crime? Maybe, maybe not. But all our stereotypes and our prejudices will push us to infer something negative about those immigrants, who are the source of all the bad things what happen in our country.

We can see how Paul Watzlawick was right when he said that in any act of communication there is a content and a relation, and that the relation is classifying the content, thus it is a metacommunication. It is quite the same finding as the one of the British scholars working in the field of the systemic functional grammar, some time later, when they were talking of the metafunctional organization of the language (Halliday, 1977). For Halliday, a social situation is a "semiotic structure" which can be described using several elements:

1. A field of social action: something that is actually going on, and that has attached a meaning within the social system;

2. A tenor of role structure: the participant relationships; as in the theory of van Noppen, they include the speech roles, those that come into being through the exchange of verbal meanings;

3. A mode of symbolic organization: the particular status that is assigned to the text (we would say discourse!) within the situation.

Language will always convey more than the information or the intention of the speaker. The text will always remain as an evidence of the meaning of the discourse in the particular situation where it took place. Confirmations of this view can come from anywhere, and in any form. Let’s look to what Jay Lemke says:

"I hope that in all our work on how meanings are made with text we will remember that the text is a product and a record of meaning making processes which are essentially dynamic. These processes are social semiotic practices, the signifying practices of a community. It is these practices that make texts and make sense of texts, dynamically,
dramatically, moment to moment, word by word, enacting meaning by words, in
moments whose meanings the words make and change."

Parts of the text structure will reveal to the eye of the analyst even the most secret
thoughts of the author. It’s frustrating to see how difficult this can be to express oneself
in words, but it can be even more frustrating to see how much more the interlocutor can
infer from what you say.

5. Conclusion

What did I mean by saying this (Lemke, 1991)? Exactly what I said! No inference
required! Can I rely on you for this? Because I cannot rely on the language.

References

Halliday, M.A.K. (1977): Text as Semantic Choice in Social Contexts in Grammars and
Descriptions (Studies in Text Theory and Text Analysis). Eds. Teun A. van Dijk and

(Ed.), Functional and Systemic Linguistics: Approaches and Uses, Berlin:
Mouton/deGruyter (Trends in Linguistics: Studies and Monographs 55).

Marcus, S (2006): Paradigme universale II. Pornind de la un zâmbet, București, Paralela
45.

Part II, Bruxelles, Presse Universitaire de Bruxelles.

Pinter.

---

2 Jay L. Lemke, "Text production and Dynamic Text Semantics", in In E. Ventola (Ed.), Functional and