A Multimodal evaluation of Malala Yousafzai's speech at Harvard University

Cristina Georgiana Ivascu
cristinageorgiana.ivascu@uji.es

Silje Handeland
silje.handeland@uji.es
I. Abstract

Through language speakers express thoughts, experiences, feelings, values and attitudes. Nevertheless, language is not only verbal communication, as multiple devices are included in interaction in order to make something coherent. Thus, people inform others about feelings through a combination of verbal and non-verbal interactions. Language is not made up exclusively of words, phrases and sentences but also of images as it is the main resource for conveying meaning. Non-verbal behaviour covers all forms of non-spoken human conduct possessing the capacity to construct communicative messages. Hence, the nature of the connection between speech and gestures has become a popular topic to study among researchers in the field of linguistics among others.

This paper presents a multimodal evaluation of an academic speech performed at Sanders Theater, Harvard University, September 27th, 2013, by Malala Yousafzai, the Nobel Peace Prize winner 2014. Even though the speech is a monologue, the speaker achieves interaction and engagement with the audience by means of using both verbal expressions and non-linguistic resources throughout her presentation. This study exhibits an evaluation of how non-linguistic resources such as paralanguage and kinetics are used as complementary tools in spoken discourse.

Key words: multimodal discourse analysis, public speech, kinetics analysis, paralanguage analysis

II. Introduction

A growing interest in the role language plays in constructing the reality that surrounds us has been recently perceived as a result of the development of new theories and methods for analysing the use we make of the language and the role it performs in social contexts (Bhatia, V. et al. 2008). Considered by several researchers (Stubbs, M. 1983; Aijimer, K. & A. B. Stenstrom, 2004) to be one of the broadest but least clearly defined areas in applied linguistics, Discourse Analysis is a field that develops very fast. It has become an object of research in numerous academic disciplines different from one another (Schiffrin, D. et al., 2001). Harris (1952) was the first to mention the term Discourse Analysis and he made the observation that particular situations require distinctive uses of the language.

There are some scholars who describe it as the study of the language in use (Fasold, R. 1990; Blommaert, J. 2005; Gee, J. P. 2014) while others define it as the analysis of the language beyond the level of the sentence. Jaworski, A. & N. Coupland (1999) have produced ten definitions for this concept although according to Schiffrin, D. et al. (2001) these ten definitions can be mainly classified into three categories “(1) anything beyond the sentence, (2) language use, and (3) a broader range of social practice that includes non-linguistic and non-specific instances of language”. This categorisation shows that discourse can be analysed from
different perspectives: going from a textual view to a critical approach of discourse as a ‘social practice’ (Fairclough, N. 1992). The definitions provided above imply that discourse should not be reduced exclusively to language.

At present, multimodality has become a characteristic feature of texts belonging to various domains in communication due to the fact that spoken as well as written language basically take place in interactions amid other semiotic modes such as: gesture, image, colour, texture, shape or spatial layout and configuration. The term was formulated upon Halliday’s (1978; 1985) view of language as social semiotic and was developed to shape “the meaning of words, sounds and images as sets of interrelated systems and structures” (O’Halloran, K. L. 2011). Despite the fact that multimodality has always been omnipresent in most of the communicative environments in which human beings are involved, they have been overlooked for a long time. However, over the last decade, as a result of the recent developments in the ‘new’ media (such as computers and the Internet) a great number of scholars became aware that the analysis of language alone is not enough to comprehend the patterns of communication (Ventola, E. et al., 2004).

Multimodal studies comprise two main fields: multimodality in language and language systems, and multimodality in other systems. The former focuses on interaction within two different perspectives: human-to-human interaction (Norris, S. 2004) and human-machine interaction (Roope, R. 1999). Norris, S. (2004) states in her work that all interactions are multimodal; interactions consist of multiple modes of communication in terms of disciplines: linguistics, sociology, education, anthropology and psychology. To achieve a complete understanding of spoken discourse, it is essential to study both linguistic and non-linguistic characteristics. The theory of multimodal communication could be defined as the use of several semiotic modes in the design of a semiotic product or event within a particular cooperation of the modes in order to express meaning (Kress, G. & T. Van Leeuwen, 2001). Therefore, language could be characterized to be only one communicative channel among many. Furthermore, a message could be defined as a combination of modes working together simultaneously, and it is therefore essential to mutually analyse both linguistics and non-linguistic features, otherwise it would be impossible to accomplish a complete understanding of spoken discourse. As a result, a Multimodal Discourse Analysis is a more complete analysis of spoken discourse (Baldry, A. & P. J. Thibault, 2006).

III. Methodology

The study exhibited in this article has followed a methodological framework for analysis applied in previous Multimodal Discourse Analysis studies on spoken academic discourse (Querol-Julián, M 2011; Querol-Julián, M. & I. Fortanet-Gómez, 2013; Fortanet-Gómez, I. & M. N. Ruiz-Madrid, 2014). The objective for this study is to analyse in a multimodal
manner the speech performed by Malala Yousafzai at Sanders Theater, September 27th, 2013, when she received the Humanitarian of the Year Award from the Harvard Foundation (available online at https://www.youtube.com/watch?v=e1tOe4SKbLU). In this study, we will simultaneously analyse non-linguistic resources (gestures, head movements and gaze) and paralanguage (loudness, syllabic duration, pauses, and laughter from the audience), as well as an evaluation of semantic resources and attitude according to Martin, J. R. & P. White (2005).

This paper includes a multimodal discourse analysis of the speech Malala Yousafzai performed at Harvard University in 2013, upon receiving the Humanitarian of the Year award from the Harvard Foundation (The Harvard Foundation for Intercultural and Race Relations, 2013). Malala Yousafzai is a seventeen years old young woman from Pakistan, who is struggling “against the suppression of children and young people and for the right of all children to education” (The Norwegian Nobel Committee, 2014), and this year she received the Nobel Peace Prize together with Kailash Satyarthi from India for their honourable work. Malala Yousafzai is a women and children’s activist, and she is at present the youngest person ever to receive both awards: the Humanitarian of the Year award and the Nobel Peace Prize. The duration of her speech is 10 minutes and 35 seconds.

In order to analyse the speech, there were some issues to take into consideration. In the first place, Malala Yousafzai’s speech is a prepared monologue; hence it is not a spontaneous spoken discourse. Another issue that we have taken into account is the position of the speaker. During the whole speech the speaker is standing behind a lectern stand, which probably can limit some of her gestures. Furthermore, the transcript of the speech was not of a professional existence, so we had to rewrite it according to the audio of the speech.

In order to achieve a more complete analysis of the speech, we have used ELAN (EUDICO Linguistic Annotator), (available online at http://www.lat-mpi.eu/tools/elan/download), which is a free tool for multimodal annotation and transcription developed at the Max Planck Institute for Psycholinguistics (MIP) (Nijmegen, The Netherlands). This tool has made it possible to design and create layers and tiers for the transcription and the annotations. Thus, it has facilitated the process of synchronizing the elements of the analysis. First, we synchronized the written transcript of the speech with the audio and the video. Then, we annotated the semantic resources and the tags in order to identify the non-linguistic resources. At the moment all the data was introduced, we commenced our multimodal discourse analysis of the speech. Figure 1 illustrates an instance of the tool used in this analysis.
Cristina Georgiana and Silje Handeland. A Multimodal evaluation of Malala Yousafzai’s speech at Harvard University

Figure 1. Sample view of multimodal annotation and transcription in ELAN

We have analysed the speech from a multimodal perspective by means of combining elements such as intonation, gestures and facial expressions, as well as the analysis of the transcript and the video recording through a systematic and synchronized procedure in order to achieve a more complete discourse analysis according to Kress, G. & T. Van Leeuwen (2001). Our research questions were whether Malala Yousafzai would use any techniques in order to capture the audience’s attention in terms of special gestures or language. Furthermore, we assumed that there would be verbal and non-verbal expressions due to her cultural existence.

IV. Results

4.1. Generic structure of the speech

The generic structure presented in the speech is the following: 1. Salutation 2. Introduction of special guests 3. Appreciation 4. The occasion 5. Personal experiences 6. The importance of Harvard University 7. Present struggle 8. Solution and hope 9. Appreciation. The examples in this paper have been selected conscientiously in order to comprise all the sections of Malala Yousafzai’s speech.

Malala Yousafzai initiates her speech with a greeting in Arabic followed by a translation into English; “Bismillah-ir-Rahman ir-Rahim, In the name of God, the most beneficent, the most merciful”. This could probably be in order to connect the language with her religion, Islam, as well as to connect with the audience. Then she continues her speech by introducing special guests, their professional occupation and the occasion of the speech; “Honorable Doctor Allen Counter, the director of the Harvard foundation,
and I guess that the president, respected president Drew Faust who I proudly say is a fast woman, founder and president of the Harvard University.” In addition she thanks the Harvard University and the Foundation, as well as the honourable guests she mentioned previously in her opening, whereas she specially shows her gratefulness towards Dr. Counter. Furthermore, she shows her appreciation in order to be with the whole audience, which she refers to as “my dear friends”. Moreover, she emphasizes the importance of the occasion. Then she moves to talk about her personal experience and she again thanks the doctor, as well as people’s prayers and God, for saving her life.

The speaker continues talking about her country of origin, Swat in the Northwest of Pakistan. She informs the audience that Swat was a target of terrorism in 2009 and that it was a horrific situation against girls and women. However, she underlines that some of them were not afraid and that they raised their powerful voices for the right of education and peace in their country. Then speaker informs the audience that Swat, although it is not a peaceful place, has been through an evolution, and that schools are reopened whereas girls have returned. The speaker refers to the audience by saying “[...] dear sisters and brothers” and tells them that they must be proud to be alumni and a part of the Harvard University. She follows by saying that the Harvard University represents great values and traditions, and that the institution has enlightened in 376 years.

Malala Yousafzai persists, by including all the people in the Sanders Theater, that all should dream about education and peace. In addition, Malala Yousafzai communicates that we all need to work for children in suffering countries and issues such as child trafficking, and child labour that children in these countries are facing. She specially mentions Syria, whereas many children are homeless and deprived from education. The speaker continues by mentioning the children of Pakistan and Afghanistan, who are victims of terrorism, and the children of India, who are suffering from child labour. Furthermore, she notifies and tells the audience not to forget that girls in countries such as Nigeria are suffering from early forced marriages and are victims of sexual violence. She tells us not to forget that in many African countries, children have not access to food or clean water. The speaker continues by emphasising the reluctance of women in several countries. Then, she gives the solution, which she believes and suggests is education. She stresses that “we can fight wars through dialogue and education”, and in order to achieve peace we should “[...] instead of sending guns, send books. Instead of sending tanks, send pens. Instead of sending soldiers, send teachers.” The speaker follows by saying that even “one book, one pen, one child and one teacher can change the world.” To end with she states her dream, and she invites the audience to dream with her, to dream that every child can have the possibility to attend school and that women’s right becomes accepted, and that we achieve equality and justice. She emphasises that in order to achieve these, we have to fight and to make “today’s dreams, tomorrow’s reality” to reach a bright future. The
speaker ends her speech by showing her gratefulness once again, before she takes a step backwards from the lectern.

4.2. Verbal and non-verbal behaviour

All through history, clothing has had a non-verbal significance; therefore it is highly important to highlight this aspect in the lecture as it transmits a powerful message and essence of Malala Yousafzai’s policy. The speaker wears a pink duphata (typical Pakistani women’s scarf) and the colour was not chosen randomly as in the Pakistani culture pink symbolizes feminine power. Therefore, this colour represents the identity of Malala Yousafzai as a women’s rights activist.

Malala begins her speech with a phrase in Arabic which she translates into English:

(1) Bismillah-ir-Rahman ir-Rahim, In the name of God, the most beneficent, the most merciful

![Figure 2. Gaze from the left to the right](image)

This sentence is said to contain the true essence of the entire Qur’an as well as the true essence of all religions, therefore she reveals the culture and religion she belongs to as well as trying to create a relationship with the audience. In addition, she emphasizes her belief and faith in God. Furthermore, by speaking in another language, which probably is not familiar with the major part of the audience, she also achieves to draw the audience’s attention towards the stage. With regards to the employment of kinesics and paralanguage, the speaker draws the attention towards herself by raising her voice at the beginning of the sentence. She does this at the same time as she looks at the audience from the left to the right as observed in Fig. 2, while by the end of the sentence she lowers her voice and bows her head.
As illustrated in Fig. 3, when introducing the name of the guests attending her speech, Malala raises her voice and employs deictic gestures at the same time as she moves her head towards her right and looks them in the eyes. Straightaway she looks at the audience while communicating and emphasizing the important positions these guests have at the Harvard University and Harvard foundation.

(2) Honorable Doctor Allen Counter, the director of the Harvard foundation, and (short pause) I guess that the president, respected president Drew Faust who I proudly say is a fast woman, founder and president of the Harvard University.

Figure 3. Hand movements co-expressed with “Doctor Allan Counter” and “Drew Fauster”

Before turning to the main issue of her speech, that is, the call for equal education rights which won her Harvard Foundation’s Humanitarian of the Year Award, Malala Yousafzai narrates the circumstances leading to her being shot by the Taliban and her uncle’s surgical intervention which saved her life.

(3) I believe...I believe that God saved my life. People’s prayers saved my life, And of course, Colonel Junaid saved my life.
When pronouncing the words “I believe”, the speaker uses a head nods and gazes at the audience from the left to the right to assert and insist in her following declaration as if trying to see if they support her statement. This instance of language co-expressed with kinesics is depicted in Fig. 4. The words “God” and “prayers” are uttered in a louder voice than the rest of the sentence to bring a focus on whom and what saved her life together with her uncle. A deictic gesture together with a head movement to the left accompanies the speaker’s presentation of her uncle. At the same time as she presents her uncle and indicates this position, the audience applauds him for the important role he carried out in saving his niece’s life.

(4) We are not here to make a long list of issues we are facing. Rather, we are here to find a solution (pause) and the solution is simple: education, education, education.

Figure 4. Hand gesture co-expressed with “I believe”

Figure 5. Gaze co-expressed with “we”
In the previous extract taken from the speech, Malala Yousafzai offers her answer for the problems existent in developing countries with regards to the lack of women’s rights. The pronoun “we” is uttered by the speaker in co-occurrence with a lateral sweep and a look at the audience as it encompasses inclusivity as exemplified in Fig. 5. Therefore, by means of using this pronoun she demonstrates to be sure that they are sharing the same interest in the topic. The word “rather” is intensified and takes place in association with up-and-down head movements, beat gestures, lowered eyebrows to express affirmation. The speaker employs pauses all throughout her speech, but she always does this with a purpose in mind. The pause is a key concept in oral communication. It is one of the main strategies speakers employ to emphasize the information they wish to deliver. The pauses previous and afterwards the unit of information help the listener to recall it better (Fortanet, I. 2008: 73). With this idea in mind, we can observe how the speaker utilizes the pause to emphasize the solution she gives for the issues commented before, a solution she considers to be crucial, and that is the need of education. The speaker uses an iconic gesture to illustrate each of the three times she utters the word “education” at the same time as she looks at her listeners and raises the tone of her voice.

In the last passage analysed which is also a famous quote by the speaker, Malala addresses world powers and asserts that if they want to see peace in Syria, Pakistan and Afghanistan they should employ not force or weapons but rather dialogue and education.

(5) If you want to end war, then instead of sending guns, send books. Instead of sending tanks, send pens. Instead of sending soldiers, send teachers.
In this sentence which contains a strong antithesis, the word “send” is pronounced in a higher volume than the rest of the words in the sentence and co-occurs with beat gestures, lowered eyebrows and gaze towards the audience to bestow authority to her request. Loudness, along with pitch, is one of the main means used by speakers to confer a special meaningful effect to words (Poyatos, F. 2002). Fig. 6 shows how beat gestures are used along with the paralinguistic feature mentioned above to mark the rhythm and intonation of the speech. Once the request is carried out, the audience answers the speaker by means of applauding her.

V. Discussion and conclusions

The findings in this paper are in line with previous research found in the theoretical framework of this study. The results exhibit the importance of paralanguage and kinesic features in order to establish a wider comprehension of spoken discourse. By analysing the speech from a multimodal perspective, this study presents a more complete meaning of the nature of spoken discourse.

The speaker follows a generic structure similar to other conference speeches as it commences with a salutation, introduction of special guests, appreciation and the occasion of the speech. Furthermore, the speaker talks about her personal experiences in order to transmit her own feeling, values, and attitude through verbal and non-verbal resources. Thus, she presents the purpose of the topics she introduces in her speech in a meaningful manner. Additionally, she provides the audience with solutions and hope in the future. At the end of the speech she once again shows her appreciation. This multimodal evaluation illustrates the practice of multiple semiotic modes in means of a semiotic product or event in terms of a particular cooperation of the modes in order to express meaning as argued by Kress, G. & T. Van Leeuwen (2001) in their theory of multimodal communication.

Likewise previous research, this study displays how a message should be defined as a combination of modes working together simultaneously by means of both verbal and non-verbal resources. According to the results, the most frequent kinesics features performed by the speaker in co-occurrence with speech where iconic and beat gestures. Argued by McNeill, D. & E. Levy (1982), iconic gestures are closely related to the semantic content of the discourse being delivered. They illustrate both by means of the form and manner of performance features of the scene displayed in the speech. Beat gestures do not possess a distinctive meaning as they are rather biphasic movements taking place wherever the hands are to be found. With regards to head movements, the analysis revealed that the most common types were head nods in association with gaze directed towards the audience. Maynard, S. (1987) observed that speakers employ head nods to mark a clause boundary, to affirm or to emphasise a point in the speech.
The following table exhibits the semantic resources of linguistic evaluation co-occurring with non-linguistic features employed by the speaker in the instances analyzed.

**Table 1:** Speaker’s expression of linguistic evaluation in association with kinesics and paralinguistic features

<table>
<thead>
<tr>
<th>Semantic resource</th>
<th>Evaluative cue</th>
<th>Paralanguage</th>
<th>Kinesics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attitude: + affect</td>
<td>(2) who I <strong>proudly</strong> say is a fast woman</td>
<td>-</td>
<td>Eye contact</td>
</tr>
<tr>
<td>Attitude: +appreciation</td>
<td>(2) Drew Faust who I proudly say is a <strong>fast</strong> woman</td>
<td>-</td>
<td>Head nod</td>
</tr>
<tr>
<td>Graduation: - intensification</td>
<td>(4) <strong>Rather</strong>, we are here to find a solution</td>
<td>-</td>
<td>Iconic gesture</td>
</tr>
<tr>
<td>Engagement: proclaim</td>
<td>(3) And <strong>of course</strong>, Colonel Junaid saved my life</td>
<td>-</td>
<td>Beat gesture</td>
</tr>
<tr>
<td>Engagement: entertain</td>
<td>(3) <strong>I believe...I believe that</strong> God saved my life</td>
<td>+Volume</td>
<td>Head nod and eye contact</td>
</tr>
<tr>
<td>Graduation: +intensification</td>
<td>(4) and the solution is simple: <strong>education,</strong> <strong>education,</strong> <strong>education</strong></td>
<td>Long syllabic duration</td>
<td>Iconic gesture</td>
</tr>
<tr>
<td>Attitude: judgement</td>
<td>(4) to find a solution (pause) and the solution <strong>is simple</strong></td>
<td>+Volume</td>
<td>Beat gesture</td>
</tr>
<tr>
<td>Attitude: judgement</td>
<td>(4) <strong>We are not here to</strong> make a long list of</td>
<td>-</td>
<td>Eye contact</td>
</tr>
<tr>
<td>Attitude: +appreciation</td>
<td>(2) <strong>Honorable</strong> Doctor Allen Counter</td>
<td>-</td>
<td>Beat gesture and lowered head</td>
</tr>
<tr>
<td>Graduation: +intensification</td>
<td>(1) God, the <strong>most</strong> beneficent, the <strong>most</strong> merciful</td>
<td>Rising intonation</td>
<td>-</td>
</tr>
</tbody>
</table>

In the process of analysing there have been some additional issues and limitations to take into consideration. Firstly, as the speech was a prepared monologue it was not of a spontaneous spoken nature. Thus, it was difficult to know whether some of the features were of a natural origin. Furthermore, another issue that we had to take into account was the position of the speaker. Since the speaker was standing behind a lectern during the whole speech it probably limited some movements. Additionally, the transcript of the speech was not of a professional existence, hence it has been rewritten according to the audio of the speech.

The findings in this study suggest that future research could focus on a multimodal analysis of a wider corpus comprising other speech given by
the speaker in different contexts in order to compare the possible similarities and differences existent in her non-verbal behaviour in public lectures. Moreover, it might result interesting to include within this same corpus other speeches of the same nature given by speakers of different native origins.

Furthermore, to speak in public might result challenging for many people, especially novice researchers and speakers delivering a speech in a foreign language. This new era of globalization and internationalization has brought about a need to present the results of one’s research not only by means of published research articles but also through oral presentations in conference presentations and symposiums (Fortanet, I., 2008). Furthermore, non-linguistic elements are highly important in these types of academic events as a competent transmission of the information depends on the selection and use of these features (Räisänen, C. & Fortanet, I. 2006). The results revealed in this study might serve to equip novice researchers with some basic information they need to take into account before delivering an oral presentation such as for example: how to engage with the audience and catch its attention by means of language co-expressed with kinesics and paralanguage and how to introduce the topic of the speech.

VI. Acknowledgements

We would like to express our very great appreciation to Dr. Inmaculada Fortanet Gómez and Dr. Miguel Francisco Ruiz Garrido for their help and guidance during the writing period. Their willingness to give their time has been very much appreciated and of great value. Furthermore, we would like to thank GRAPE (Group for Research on Academic and Professional English) for their collaboration and information regarding this topic. This research is part of the project funded by the Ministry of Science and Innovation of Spain. (ref. FFI2011-24269).

VII. References


Cristina Georgiana and Silje Handeland. A Multimodal evaluation of Malala Yousafzai’s speech at Harvard University


The Harvard Foundation for Intercultural and Race Relations (September 27, 2013). Available Online:

