

CHAPTER 1 WHAT'S CHIC?

“Reality exists outside language, but it is constantly mediated by and through language: and what we can know and say has to be produced in and through discourse” ~ Stuart Hall, *Encoding/Decoding*

“Chick,” I said. “It means a baby chicken - or it’s a slang word for a girl.”

“Or something with fashion?” my student said.

“What...?!?”

My young student, probably about 10 years old in 1997, had somehow learned a strange low frequency word. After some further explanation, I figured out that he was referring to ‘chic’. He’d read this word in the teletext prompt while watching American television shows in his home in Vaslui, Romania. There weren’t an abundance of shows in English to choose from, but *Seinfeld* had started about a month before, *Baywatch* had been on for awhile, and I knew from previous questions that he liked to watch television.

He, and many of my other students, had an envious proficiency and command of the English language in just their third year of English classes. All of their teachers were trained in British English, and had British accents. But I had many students who had American accents, yet had never met an American before my arrival. They were television and movie junkies as well.

I started to think about simultaneous reading, hearing and watching as learning tools for language learning. Upon my return to the United States, I started researching

the use of closed captioning as a tool in learning and in retention of vocabulary. This dissertation is an expansion of my earlier attempts at understanding this combination of input; I've now branched into the interdisciplinary studies involved in exploring multimodality.

1.1 Multiple Multimodalities

The term multimodality is used in different fields of language study and different disciplines to refer to quite different concepts. It can refer to the physical modalities such as seeing, hearing, touching, tasting, etc. This is generally the term used in psycholinguistics and cognitive science, for example when discussing synaesthesia studies or pedagogical uses of technology (e.g. Paivio, 1986; Mayer, 2001; Nelson, 2006). Somewhat related¹, *multimodality* can also be used when talking about analyzing conversational discourse, such as the use of pragmatic cues or gestures to communicate (Norris, 2004). *Multimodality* is also used to refer to the combination of different semiotic qualities of a physical representation, such as the use of color, font size or pictures in a magazine article or a children's drawing (e.g. Kress & Van Leeuwen, 2001). The common theme between all of these uses is the premise that communication is being conveyed through more than one 'channel', or more than one signal. A combination of cues is used to breach the gap between the speaker and hearer, or the author and the

¹ Statistics also uses affiliated terms, e.g. '*bimodal*' for the occasion when the data does not fall into a normal curve, rather into two areas of distribution. In the medical sciences, *multimodality* is used to refer to the use of more than one medical test or more than one treatment for a diagnosis.

viewer, encoder and decoder, or whatever two terms that could be used to describe a communication event.

But to try to break a multimodal event into its pieces in order to observe and analyze its constituents is quite challenging. For one reason, meaning may not be contained in only one of the modalities; often it is a synthesis of the different channels of meaning. Kress & Van Leeuwen, in the preface to their text *Multimodal Discourse* (2001), admit to the complexity (and frustration) of trying to break down multimodal texts into their individual pieces and instead had to create a framework and concepts to go with their analysis. Sigrid Norris also needed to develop a framework to analyze the combination of modes of communication in interaction, as she looked at the use of modes such as gesture, gaze, intonation units, and even levels of awareness (2004). In order to analyze the different aspects of websites, Lisbeth Thoriacius adapted Jakobson's Communication Model to break a website into its different functions and their rhetorical appeals (2002).

However, all of these frameworks for analyzing the parts also take into account the 'whole', and that the whole text, whatever the text is, "transcends the collective contributions of its constituent parts", according to Hull & Nelson (2005), who worked with digital multimodal video texts created by teens at an urban community center (2005). Perhaps the ultimate expression of this challenge of multimodal texts and their analysis comes from an example originally used to describe language and thought by Lev Vygotsky. In an analogy of a scientist looking for answers as to why water extinguishes fire, Vygotsky illustrates that breaking this element into its pieces will not yield the

answers to the question since hydrogen burns and oxygen feeds a fire. Instead, it is the union of the parts that is the process that should be studied (1986, p. 4). It is this very notion that the multiple layers of meaning expressed in different modalities means more together than when in monomodal expressions, and that mass media can be accessed by such a large percentage of the world's population, often for free or a low cost, that led me to become interested in using media as a learning tool. The physical modes (channels) of meaning gave me reason to think that learners with different proficiencies in the different modalities (e.g. listening or reading) would be able to use their preferred modality to learn and gather as much meaning from a text as possible on an individual basis.

Up to now I have touched on a few of the issues that I encountered while trying to decipher all the different uses and definitions across disciplines, in the quixotic attempt to bring them together in one dissertation. This doesn't quite happen. One of the issues that I encounter is my adherence to a functional approach to language and the use of authentic texts, which leads to methodological entanglements. The inherent mix of variables in a multimodal presentation of a text, and the desired analysis of a variable that sits within the context of its original presentation without alterations, is fascinating and difficult. As a result, in order to gather research data about one of the modalities separately, an alteration to the original text was made, and the text in Condition 1 was reduced from its multimodal form to a monomodal form with which information about reading closed captioning can be gathered. Otherwise, all errors in the text were kept and the multimodal presentation of the story was presented with all of its channels. The use of authentic texts is not unprecedented, although it is in the minority in eye movement data

methodology. These issues of the text, the conditions, the participants and the literature surrounding multimodal texts in reading and learning are discussed further in the remaining chapters. Below, I give a preview of the research questions discussed in Chapters 5, 6, and 7 as I explore the modality of moving print and different uses of it by different participants:

1.1.1 Overview of Research Questions

READING PATTERNS OF MOVING TEXT:

- How do the reading patterns of dynamic text differ in comparison to the reading patterns of static text? And, is there a difference between native and non-native speakers of English's reading patterns of dynamic text?

VIEWING PATTERNS WHEN TEXT IS ADDED:

- In what ways do the reading patterns of dynamic text change with the addition of the multimodal environment? And, are there any similarities or differences between native and non-native speakers in the reading patterns of dynamic text?

SITUATING THE LEARNER: SMALL CASE STUDIES

- What are the relationships, if any, that can be established between the individual viewer's reading patterns and the self reported background history related to multimodal use?

1.2 Dissertation Overview

This introduction gives the impetus behind the research questions and places this dissertation in a broad context of multimodality and learning. It also presents the reader with a framework in which to place the study: second language learners using different channels of information available in a multimodal event as tools for improving and gaining comprehension. It gives the background for setting this dissertation in the context necessary to move in and out of the different disciplines revolving around this study and places the participants in a multimodal framework. Chapter 2, the literature review,

narrows the context to that which directly affects the conditions and questions of this study. Chapter 3 describes the methodology to discover the answers and the materials that were created in order test for them. Chapter 4, 5 and 6, the analyses of the three research questions, details the analysis of the data collected, breaking the large text into smaller segments, and Chapter 7 connects the questions and the answers, listing limitations of data collection and possible future research venues, including pedagogical applications gained from this study.

1.3 The context behind the texts and the participants

In the remainder of this chapter, I will establish the broader context for this research study by looking at the parts that synthesize into the whole, as mentioned above. One complication is the very interdisciplinary nature of the topic of this study. Multimodality is at the core of my research interests. I believe that it can be used to give reinforcement and options to language learners and interpreters of a text. However, different disciplines give different meanings to the same terms: ‘multimodality’ in psychology is quite different from that of literacy studies, and for that matter, from the term ‘modal’ as it is used by grammarians. A seemingly simple term ‘text’ is defined differently by scholars in reading, media, linguistics and rhetorical studies. This chapter will set up a general context or framework in which this study can be grounded and viewed within a larger picture.

However, one of the most important aspects of this chapter is to present the reader with the context for the present study. The different channels of information available during a multimedia event (explained below) are complex. In order to explain how the

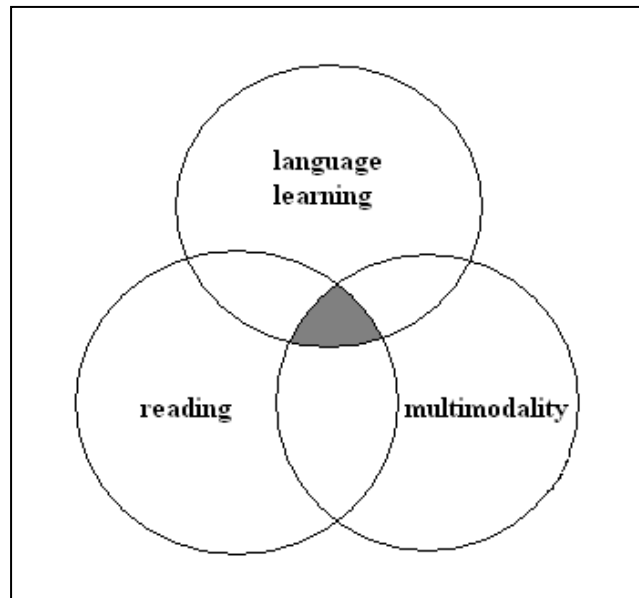
second language learner fits into the framework of a multimedia event, Chapter two looks at each factor of Jakobson's Speech Event, and complicates it by expanding it into a Multimedia Event for second language learners. Once the second language learner is thus placed, the relevancy of the study to the Multimedia Event and the availability of the different channels of information for use in interpretation by the second language learner are presented.

1.4 Merging across disciplines

This study touches on many topics – and necessarily so, as multimodality by its very definition indicates multiplicity in its essence. Multimodality is enormous in its breadth, and it will be refined in its scope and relevance to this study later in the chapter (for a definition, please see the List of Terms at the beginning of the dissertation). This study looks at a minute area of multimodality, that of reading a text with or without the additions of audio and visual texts for the purposes of supporting better comprehension for second language learners (see Figure 1.1 below). But first, the context of multimodality is needed, in order to place this study within it. Then, the study is placed within a framework of communication, the Speech Event Model, so that a common set of ideas can be used for description. Following this broader context, the focus will narrow to the theoretical backgrounds of reading and the idea of a transaction with the text including that of L2 readers, along with a review of past research in closed captioning/subtitling, learner strategies, and a brief overview of past eye movement research will finish. More specific literature regarding these will also appear in the analysis sections. In the end, this study seeks to find the cross section between language

learning (and the language learner), reading, and multimodality, as illustrated in the figure below.

Figure 1.1 Illustration of apex of study: The cross section between language learners using reading in multimodal environments



1.5 Multimodality: Different definitions in different disciplines

There is more than one use of the term ‘multimodal’ in literature. Various authors use modality to mean different literacies, or abilities to interpret and manipulate verbal, graphic, acoustic, etc, texts, including the ability to semiotically interpret, for example, different genres, the internet, or advertisements. Other forms of the word ‘modal’ include linguistic terms, such as believability or strength related to verbs (could, would, might).

1.5.1 Multimodality 1: Sensory ‘modes’:

In Psychology, a multimodal event is one in which more than one modality is used for perception. A famous example of a bimodal event is the McGurk Effect

(McGurk & McDonald, 1976), in which the sound produced (/ga/) is different than the movements made visually by the speaker's mouth (/ba/). The resulting effect is that the person who is observing this altered bimodal event perceives the sound /da/ when the two modalities, visual and aural, blend in perception. Multimodality, in this sense, refers to instances when the modalities are physical in nature and there are multiple inputs of different physical sensory modes, such as visual, aural, tactile, etc. An interesting example of an integrated cross-modal experience is synaesthesia, in which activation or use of one physical mode triggers another, e.g. hearing music and seeing colors.

In more common language experiences, recent experiments have shown the bimodal synergistic effects between visual and aural stimuli. In these studies, response times are faster when the stimuli are redundant (e.g. seeing a dog and hearing a dog barking) (cf. Giard & Peronnet, 1999; Fort et al., 2002). Experiments involving affective prosody (e.g. seeing a face while hearing a sentence, both on a happy to fearful continuum) have shown that aural stimuli can bias perception of visual facial expression. In other words, humans integrate the two modes and are influenced by both in an automatic and mandatory response, regardless of attentional resources (cf. Gelder & Vrooman, 2000; Vrooman et al., 2001; Pourtois et al., 2005). In language learning, gesture, too, has been shown to play a role in comprehension in a multimodal form of communication (cf. McNeill, 1992). Breckenridge Church, Ayman-Nolley, & Mahootian (2004) conducted a classroom experiment in which gesture either did or did not accompany instruction in English to children with an L1 in Spanish. Fifty percent of the children who received the concrete, representational gestures improved their

understanding in a post-test compared to only 20% of the children who received monomodal instruction, or without gesture. The authors point out that gesture should be used to enhance instruction for learners and for those learning in a language other than their first language.

It is evident that humans use multiple sources of information to inform themselves of the world around them; this is particularly useful in the use of media as a tool for learning a second language: using subtitling/closed captioning may reinforce the modal inputs. These types of psychology experiments, in which the sensory modalities are separated and strictly controlled, seem to go against a transactional or sociocultural view of language in that they are based on the premise of studying the unitary nature of complex items, instead of the parts. Vygotsky (1986) uses the metaphor of studying water: If its composition is studied for its extinguishing properties by separately studying oxygen, which fuels fire, and hydrogen, which burns in contact with fire (pp. 4-5), then its combination is puzzling.

Likewise, multimodality should be studied in its totality. However, by studying perceptions of bimodal stimuli, further advances in discovering how the brain functions and perceives stimuli actually provide support for the complexity of thought and language. Therefore, this study looks at both conditions: one in which the channels of information have been isolated and so only dynamic text (closed captioning) is shown, and one in which all of the modalities are present (aural, visual and print textual channels of information).

1.5.2 Multimodality 2: Modes of semiotic ‘literacies’ of different ‘modalities’

But other definitions of multimodality exist in the literature. Kress and Van Leeuwen (1995) argue for taking a fresh look at the economic changes in the post-industrial world, in which the changes are “knowledge based” and “information-driven” and in which “information of various kinds may be more aptly expressed in the visual rather than in the verbal mode,” including looking at human semiosis in the domain of communication and representation (p. 183). In the past, writing was primarily perceived as monomodal in perception and valued over speech. The senses were often thought of as separate. Kress and van Leeuwen give examples of this monomodality such as writing without illustration, paintings used canvas and the same mediums (oils), in concerts musicians dressed in similar uniforms, whereas today there is more crossing of boundaries. The authors state the need to “explore principles behind multimodal communication. [They] move away from the idea that the different modes in multimodal texts have strictly bounded and framed specialist tasks” and give an example of the prescribed modality of film in which “images may provide the action, sync sounds a sense of realism, music a layer of emotion, and so on, with the editing process supplying the ‘integration code’, the means for synchronizing the elements through a common rhythm...” (p. 2). In opposition to this, Kress and van Leeuwen see “a view of multimodality in which common semiotic principles operate in and across different modes, and in which it is therefore quite possible for music to encode action, or images to encode emotion” (p. 2) in which the ‘same’ meanings can often be expressed in different semiotic modes.

Hull and Nelson (2005) explored the semiotic use of multimodality in digital texts. In a discussion about the use of multimodality in literacy pedagogy, they state, “the point is that images, written text, music, and so on each respectively impart certain kinds of meanings more easily and naturally than others. We believe that this idea is the most crucial conceptual tool that one must bring to bear in understanding the workings and meanings of multimodal texts” (Hull & Nelson, p. 229). The authors explore the use of multimodality in different semiotic modes in digital storytelling (including spoken words, images, music, written text, and movement and transitions), finding that the different modes act synesthetically in their multimodality through four qualities: a) the visual pictorial mode can repurpose the written, linguistic mode; b) iconic and indexical images can be rendered as symbols; c) titles, iconic and indexical images and thematic movement can animate each other cooperatively; and d) modes can progressively become imbued with the associative meanings of each other (Hull & Nelson, p. 239). In the end, the authors call for a widening of the definition of writing in which multimodal composing is included (p. 252).

1.6 Placing multimodality into communication: The Speech Event

For this study, in order to place everything into a unified structure, all aspects will be grounded in the Speech Event Model as laid out by Roman Jakobson (1960). No attempt is made to offer the model as the only way to look at language learning or at the use of multimodal texts, but merely as a streamlined and effective one that gives the reader and the author a way to establish terms and relationships within this research. In fact, for this study, the Speech Event Model could really be re-titled as a Communicative

Event Model, or one that works with a synchronous, dialogic speech event occurrence. Also relevant is a discussion of Literacy Events, but this will be discussed in the next chapter. At this point, it is sufficient to attend to the broad picture: all interactions are transactions, in that something is given and something is changed. In Sociocultural Theory (Vygotsky, 1986), learning is not in isolation but rather it is the interactions between people that can further assist learning, while in Rosenblatt's Transactional Theory of Reading and Writing (1994), as well as in Goodman's Transactional Sociopsycholinguistic View (1994), the reader is involved in an interaction with the text, and brings to the text world knowledge and expectations that change as the reader reads. This dissertation revolves around multimodality by concentrating on one of its modalities: reading. In Chapter 3, a background of relevant reading literature and theories will be explored. At this point, for this chapter, the idea of an interaction is the main focus.

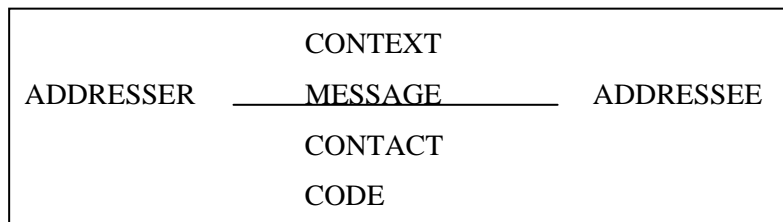
Similar to the idea of interactions and change, Jakobson discusses the Saussurian ideas of *langue* and *parole* as he furthers their definitions as *code* and *message* (1960/1990). As a speaker speaks (the message) he/she is adding to the fluidity of the code, or the larger summation of language that is used by the language's speakers. This interaction and change is important for second language learners as a language is in motion and never static. For this reason, Jakobson's Speech Event Model will be laid out so that the dynamic relationship between second language learners and the language can be analyzed.

1.6.1 Jakobson’s Speech Event Model laid out

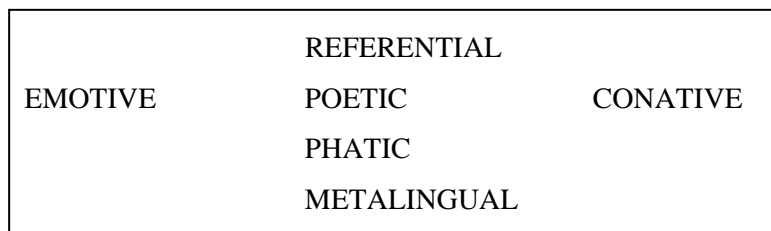
Jakobson used Karl Bühler’s tripartite model of the speech event to study language as a dialogue by expanding it in its early stages to four factors: a speaker (encoder), an addressee (a decoder), a thing referred to (context) and then he added the message (or, the instance of parole communicated from speaker to addressee) (Waugh & Monville-Burston, p. 15). Jakobson’s actual Speech Event Model, as later laid out in his plenary speech for a “Linguistics and Poetics” conference, is composed of the six factors (addresser, addressee, context, message, contact and code) which are variously focused on in the message by six functions (emotive, conative, referential, poetic, phatic and metalingual) (Jakobson, 1960, p. 72-3) (see Figure 2.2 below). It is a broad model in its deceptive simplicity, one which is easily used as a tool to analyze not only spoken but also written discourse, including one of Jakobson’s passions, poetry.

Figure 1.2 The factors and functions of the Speech Event Model (Jakobson, 1960).

Six factors of the speech event:



Six functions of the speech event:



Its ease of application stems from its ability to encompass an entire speech event within the model, and so it facilitates further expansion of the description and breakdown of communication. As the process of communication necessarily entails at least two parties, (or the essence of a dialogue whether it be inner-speech (Vygotsky, 1986) or in face-to-face interaction), Jakobson's model captures this dynamic nature of the speech event in the factor he labeled 'context'. To define the parts of the Speech Event Model, as they relate to an attempt at a successful communicative act, the speech event physically begins with the ADDRESSER, who, through CONTACT, sends a MESSAGE (constructed on the basis of the CODE), which is embedded in a CONTEXT, to the ADDRESSEE. If successful, the addressee will understand and be able to generate a communicative act back if he or she so chooses.

As an example of this, two people, George and Leah, are talking. George says, "I'm hungry – wanna catch a bite to eat?" to Leah. In this example, George is the ADDRESSER, Leah is the ADDRESSEE, the MESSAGE is the utterance, which is in English (the CODE) and the CONTEXT is that it is dinner time as the two are sitting on the couch (CONTACT). But what happens in a multimodal environment is that part of the MESSAGE could be conveyed by pragmatic means, such as indirect speech acts, or by body language in the form of gestures, which could reinforce the simple message, or could create confusion. While the message in this case seems fairly straight forward, it could be misconstrued by a second language learner for whom the CODE is not as familiar, especially if idiomatic phrases or slang terms are used. If Leah looks confused, George can re-state his question using different words or gestures or intonation, etc., in

which case he is altering his speech act for his addressee and through the interaction of the speech event the addressee can attempt understanding.

Jakobson's speech event model includes the broad nature of the context, and indeed, without the dialogic factor communication would fail. This would seem to pose a problem for using this model as the base framework for this study on multimodality and the use of multimedia, however I contend that no use of multimedia is truly a one way communication event existing in isolation. Rather, a multimedia text is always created for an audience so that a viewer will understand and this, coupled with the inner dialogue of the viewer and the use of the multimedia as a learning tool, necessarily creates the conditions needed to fill the schema of the addresser, addressee and the remaining factors and functions of a speech event. Mainstream multimedia texts are rarely created to exclude viewers, particularly multimedia texts which aim at large audiences (e.g. television, film, and the internet). Again, the functions and factors as they relate to a multimedia text and a viewer will be further explained as the chapter progresses and the discussion progresses to a sociopsycholinguistic, transactional interaction between the viewer and the text.

1.6.2 A communication event: The multiple sides of the addresser

The communication event then is composed of an interaction that can be analyzed according to its functions and its factors. To start with, the degree and strength of the orientation of the function depends on the composition and intent of the message; for example, a message focused on the context would carry a greater degree of the referential function and perhaps less of the poetic or metalingual function. In terms of a multimedia

example, this might be something similar to a newscaster (the addresser) talking about a news item (the message) for the benefit of the audience (the addressee) in which the intent is to deliver the news (the message is focused on the context, the referential function). A contrastive example may include a commercial for a bath bubble product, in which the language that is used is poetic in its alliteration and flowery language is used to convince the viewer that this bath bubble product will relax the buyer (and so the message is quite emotive in its function by the addresser, or the commercial's director, and uses poetic language and imagery in its message to reach the viewer, or the addressee who is the receiver of the conative function). The point is that a message in and of itself has many components, and so an addressee, or the person who must decode the message, must take this myriad of components into account to try to understand what the addresser meant. If the addressee is not of the same culture, one or more of the components may not be interpreted or decoded correctly.

For a second language learner, not only is the message a difficult path to navigate at times, but also the task of figuring out the context of the addresser and how the 'framing' of the context places the addresser, especially if the addresser is embedded in a multimedia text such as a movie or television show. Besides Jakobson's framework, there are other, similar, frameworks that illustrate the complexity that a person encounters when engaging in communication. One of the more prominent sociologists/anthropologists in the topic of communication and participation, Erving Goffman, divides the speaker into three distinct roles. In his participation framework, Goffman refers to the speaker, or the first person "I" in deictic terms, as the 'animator',

‘author’ and ‘principal’, who can coincide and be the same person, or can be differentiated into the one who produces the message (in voice) (the animator), one who is responsible for the selection of words (the author), and the one whose beliefs or positions are being represented (the principal) (Duranti, 1997, p. 297-8). For example, in the production of media texts, such as film and video, the animators would be the actors and the characters that they play, the authors would be the scriptwriters, while the principals would be the directors and producers.

Verschueren (1999) further expands upon the multiplicity of the addresser. In a Bakhtinian sense, the addresser, or the ‘utterer’, can consist of different layers, or voices, relating to the original ‘source’. As the “utterance is simply defined as a stretch of discourse produced by the same utterer(s), with a relatively clear beginning and end” (p. 81), the utterer producing the utterance echoes that of Jakobson’s ‘addresser’ in that the discourse can be either spoken, or written. In the case of written discourse, Verschueren designates the option of the author as the main utterer although an ‘embedded utterer’, who is one (or more) of the characters, produces the utterance. An example of the embedded utterer in spoken discourse includes mass media, such as “in the world of television, an interviewer and an interviewee are consciously engaged in the uttering and interpreting of linguistic forms which they know or want to be embedded into a wider communicative event with the television network at one end and a mass audience at the other, either simultaneously or consecutively” (Verschueren, p. 82).

This concept of the embedded utterer is important, especially in reference to the influence and use of media and technology in modern terms. As such, Verschueren lays

out the terminology for the multiplicity of voices and sources surrounding the main utterer, using utterer^E for the embedded utterer, and utterer^V for a virtual utterer, or a reference to an utterance which might or might not have been uttered. He also designates the ‘source’ as the origin of information about which one is uttering, which can again be further specified into a source that is not the actual original source but a source once removed (source⁻¹), or a virtual source (source^V) in which the actual original source of information is unknown. All of these utterers and sources, virtual or removed, may be used within the course of a speech event. For the purpose of this study, it is suffice to say that, although a communication event can be quite complicated when involving a media text, the audience including second language learners need to navigate the multiplicity of the addresser and eventually the addressee in order to interpret the message.

1.6.3 A communication event: The very necessary contact and context

Second language learners have to be able to engage in a communicative event beyond just receiving the sound wave produced by speech: there is more to interpreting the message than translation from one speech event to another. The addresser/utterer/speaker has to have some way to communicate the message, and so the factor of CONTACT is designated as, “a physical channel and psychological connection between the addresser and the addressee, enabling both of them to enter and stay in communication” (Jakobson, 1960, p. 73). The contact factor is related to a focus on the message by the phatic function, and so the purpose is not only a physical presence, but also an underlying psychological understanding of the communication act. Important for

this study and the idea of multimodality in general is the premise that a message may be conveyed semiotically through non-linguistic communication (e.g. through visual representations, through music, etc.). Therefore, in second language acquisition, a participant may not understand all of the pragmatic communication that could fall under the limits of the contact factor and its phatic functions in a communicative event but may be able to pick up some of the message through other means.

The CONTEXT is emphasized by the referential function, or the “leading task of numerous messages” (Waugh & Monville-Burston, p.15), or the communication of information. Jakobson placed messages in the contiguity of a context: “whether messages are exchanged or communication proceeds unilaterally from the addresser to the addressee, there must be some kind of contiguity between the participants of any speech event to assure the transmission of the message” (1956, p. 120). Between the participants, an internal relation must exist, through “the separation in space, and often in time...there must be a certain equivalence between the symbols used” in order for communication to be successful.

As an example of how the context, and the signs selected, must be understood by the participants in order for communication to take place, Barthes uses “The World of Wrestling” (1957) as a communication event in French (entertainment) culture. As a spectacle, rather than a sport, wrestling uses signs to signify meaning to the audience; without the context of the spectacle, and knowledge of its semiotic system, the signs would vary in their interpretation. But within the world of wrestling, the actions take on

clarity in their meaning as representations for larger concepts as an “intelligible spectacle” (Barthes, p.20).

An illustration from a multimedia² standpoint might further resonate: if a viewer starts a film or a TV series in the middle, much of the context is lost and the storyline is confusing. The beginning of the narrative is needed for full comprehension and relevancy of the current plot. In the same way, for the effective exchange of information a communication event must exist in a context that is known to both the speaker and the hearer. Likewise, in second language acquisition, the context is highly relevant to comprehension. An outsider entering into a speech event with a non-familiar context (e.g. a new community, a new job, a new class, a new culture, a new language) will encounter initial difficulty, as the semiotic clues may have different meanings. For a second language learner, the instances of *parole*, or the messages, may be slightly different from the phrases and language learned in a foreign language classroom and static textbook, and so when entering into an environment, or communicative event, the learner may not be able to correctly interpret or comprehend the message.

It is possible, however, that a modified multimedia text may be able to assist the learner in accessing more recent and relevant instances of *parole* within a situated *langue*. In the previous section, I indicated that I believe that all multimedia texts are modified to some extent in that they must appeal to a wide audience in order to be successful media texts. In the following section, I will further explain how meaning in media texts is

² Multimedia is the use of words and pictures to present material (Mayer, 2001) while multimodality refers to the broader definition in which more than one mode is used to communicate or present material (e.g. aurally, haptically, visually, etc). See earlier in the chapter for definitions and uses of the term multimodality.

shared between interlocutors, and further explain the idea of accessing the *langue* and *parole*, and a modified multimedia test in the terms of *recipient design*.

1.6.4 A communication event: Message & code

Language has fluidity. The message and the code exemplify this. Jakobson defines the MESSAGE by illustrating its essence and integration with the code, and so the two are intertwined. The MESSAGE and the CODE are also known as *langue* and *parole*; the instance of *parole* includes the influence of *langue*, and the necessary semiotic knowledge in order to interpret the instance. In other words, the utterance by an individual (*parole*) is influenced and incorporated within the intersubjective code of society (*langue*). In “*Langue and parole: code and message*”, Jakobson discusses, in depth, Saussure’s definitions of *langue* and *parole* by contrasting them with his own extension of the two aspects of language. Whereas Saussure envisioned the two as distinctly different, Jakobson found integration instead of dichotomy.

Saussure argued for linguistics to concentrate on *langue*, as *parole* was confusing because of its individuality (Culler, p. 41). In this framework, Saussure worked towards an explanation of, what he called, the arbitrary nature of signs and in so doing needed to separate the social nature of *langue* from the individual nature of *parole* and so create a controlled system to study. Jakobson did not agree with this strict dichotomy, as he saw *langue* and *parole* both at once as social and individual in their aspects. One of his main arguments against the notion of *parole* as an individual act is based on the idea of language as a dialogic action, always containing a speaker and listener (even if the listener is internal), and so because “a conversation is a whole of which each remark is

but a part that cannot be separated from the whole except in an artificial manner...

[parole] is an intersubjective phenomenon, and consequently, a social one” (Jakobson, 1984, p. 93). This relationship between langue and parole as individual and social in nature is important for second language learners: as access into a language community progresses, a second language learner will hopefully learn more about the intersubjectively held meanings and be able to use them as they engage with other members. One way of accessing a new language community is through the multimedia produced by its members, for its members.

The message, or parole, is also something that must be decoded by the addressee in order for the intention of the communication act to be successful. In fact, Jakobson later altered the terms, addresser and addressee, to encoder and decoder (Waugh & Monville-Burston, p. 489). The concept of encoding a message which must be decoded, using knowledge of the code and message, is echoed in the process of the ‘recipient design’, or the “process of adapting forms of expression to interpreter roles” (Verschueren, 86). In this process, “speakers design their speech according to their ongoing evaluation of their recipient as a member of a particular group or class” (Duranti, 299). In recipient design, the utterance (or the message) is “designed specifically for an intended audience, to ensure continued attention as well as the desirable level of understanding” (Verscheuren, 86). If it is face-to-face conversation then the design, encoding, and enactment of the message is simultaneous, while if it is a communication event that is pre-drafted, pre-recorded, or rehearsed, such as in a media event, the audience is considered and appropriate language use is judged ahead of time.

Recipient design in the use of the message (parole) within a common code (langue) is relevant to another of Jakobson's topics of inquiry: that of the axis of equivalence. Jakobson emphasized the selection and combination of units of language when building utterances, starting with lexical items and grammatical structure combined to form sentences and on up. But these choices are limited in that the utterances must be built "from the lexical storehouse" common to the addresser and addressee (Jakobson, 1953, p. 117), and thus "the efficiency of a speech event demands the use of a common code by its participants" (p. 120). The message, then, is composed of these choices which fall into two modes of arrangement: combination and selection.

On the level of the building of an utterance (for Jakobson uses these two modes in the selection and combination of units as small as phonemes as well), the "addressee perceives that the given utterance (message) is in a combination of constituent parts (sentences, words, phonemes) *selected* from the repository of all constituent parts (the code)" (p. 119-120). The combination of words follows a similar path set by Saussure, in the paradigmatic and syntagmatic choices. It is within this dichotomy that the addresser chooses the message. It is also within this dichotomy that the addresser must choose linguistic signs that the addressee can decode. Within the ideas of accommodation and recipient design, then, the choices for a message are restricted or expanded depending on the common knowledge of a similar code, as well as knowledge of the recipient. In second language acquisition, this choice of signs is fairly crucial; in media the rhetorical devices chosen are likewise crucial to the intended decoding by the recipients. Therefore, the specific codes chosen for accessible decoding by a wide

audience also has the potential for decoding by the extended audience: those viewers, such as second language learners, who are outside the intended language community but are still interested in accessing the multimedia event.

While the message is obviously an important part in a communication event, for second language learners the code, e.g. the choice of which language is to be used in a conversation, is crucial. The code must be at least a partial commonality between the addresser and addressee in order for successful communication to occur. Much of the code is agreed upon by the society in which the code rests. It is, as Duranti describes it, “commonality between the participants, reference to cultural occurrences and events, common expectation for cultural speech acts” (pp. 318-9). The addressee therefore must be able to select the correct meaning that the addresser has embedded in the utterance.

1.6.5 A communicative event: Genre & intersubjectivity

Oftentimes the meaning is embedded in a selected common ‘genre’ or a typical form of language, or utterance, including certain typical kinds of expressions used within genres, i.e. greetings, farewells, table conversations, narrations, or as Bakhtin writes, “typical situations of speech communication, typical themes, and, consequently, also ... particular contacts between the *meanings* of words and actual concrete reality under certain typical circumstances” (Wertchsh, p. 60-1). Bakhtin referred to the existence of different types of genre in speech as a necessary element; to further explain its integral part of deciphering the code, he states that “a speech genre is not a form of language, but a typical form [a type] of utterance; as such the genre also includes a certain typical kind of expression that inheres in it” (quoted in Wertsch, p. 61). Wertchsh writes that, in

Bakhtin's view, "it is no more possible to produce an utterance without using some speech genre than it is possible to produce an utterance without using some natural language, such as English" (p. 61).

If the genre is so integrated with the code of the society, then the outsider, or a second language learner, could encounter difficulties decoding the addresser's selection or use of genre and its multivoicedness, or 'ventriloquism'. Indeed, Bakhtin addresses the issue of necessary knowledge of the code when he discusses the multivoicedness of speech genres in writing that "the expressiveness of individual words is not inherent in the words themselves as units of language, nor does it issue directly from the meaning of these words: it is either a typical generic expression or it is an echo of another's individual expression, which makes the word, as it were, representative of another's whole utterance from a particular evaluative position..." (Bakhtin, p. 85). These expressions of words within different genres are used within pragmatic contexts of speech communication held intersubjectively by a speech community, one of which second language learners may or may not be aware.

On the other hand, the outsider or second language learner could also use the recognizable patterns known genres to aid in comprehension, such as those found in media texts. A television show about the news is supposed to adhere to certain features such as an unbiased truth in reporting, although there are varying degrees of reliability and veracity within this genre. However, the point is that there are similar features across different news shows that are recognizable and characteristic of the genre, such as the reporter or the newscaster, the presentation style on the screen of a person speaking with

the presentation of additional pictures or video clips that are related to the newscast.

These characteristics make the genre easily recognizable and also somewhat predictable in its presentation of information. Likewise, many films follow a similar narrative pattern, with introductions, complications and a climax in the plot near the resolution at the end.

The idea of the intersubjectivity of the code is also described by Eco as he integrates semiotics within the code when he states that “any linguistic innovation can work only when accepted and integrated by social consensus, and the same happens with the other communicative systems,” and adds that, “any semiotic system is submitted to general semiotic laws and functions as a code; but such codes are also linked to specific communities (from village to ethnic unit) in the same way in which a language produces its subcodes linked to given professions or activities...” (Eco, p. 112). The code, or language, including the use of speech genres, held intersubjectively by a society describes the necessary semiotic knowledge and commonality needed by the addressee to decode a message, and one that must be accessed or partially shared by the second language learners in order to effectively understand the communicative event.

1.6.6 A communication event: The all-important addressee

Finally, as the last of the six functions of Jakobson’s Speech Event, the ADDRESSEE can be discussed. In this study, the addressee is very important, as this Speech Event function is comprised of the participants. The addressee, or the recipient, decoder, interpreter, or hearer, of the speech event needs to have knowledge and understanding of the context and code of the message in order to interpret the message.

This involves a similar “equivalence between the symbols used the addresser and those known and interpreted by the addressee (Jakobson, 1954, p. 120). For Bakhtin, the addressee “can be an immediate participant-interlocutor in an everyday dialogue,...a more or less differentiated public, ethnic group, contemporaries, like-minded people, opponents and enemies, a subordinate, a superior, ... it can also be an indefinite, unconcretized *other*” (p. 87). In other words, the addressee echoes the complexity of the addresser. It is within the addressee that the communication event is successful or not and, as Verschueren stated it, “once spoken, an utterance is no longer in the control of the utterer: it begins to lead a life of its own in the mental worlds of others” (p. 169).

Within his production format of an utterance, Goffman separates the addressees, or hearers, into two groups: ratified or unratified participants. These distinctions are important as “speakers will modify how they speak, if not what they say” (Goffman, p.136) depending on the audience and the participants. In a similar vein to Goffman, Verschueren complicates the interpreter much like he does with the utterer. In order to distinguish the multiple audiences possible of an utterance, Verschueren labels those interlocutors closest to the utterer, or those directly involved in the speech event, as direct addressees, while those who are also listening and acknowledged participants, or presences, are labeled as side participants. All other are non-participants, but since these non-interpreters are within hearing range of the communication event, they are also factors accounted for and labeled as such. Bystanders are those who are nearby, overhearers are possible interpreters and these include listener-ins (those that are attempting to listen) and eavesdroppers (those that are secretly listening) (Verschueren, p.

83). An addressee can even be an intermediary, or someone who is told a message with directions to pass the information on to another, such as a press release given to a journalist who gives it to be printed or released through a media outlet to a wider population (a ‘removed addressee’, or addressee¹)(p. 85). An important note regarding the addressee and media is discussed by Stuart Hall in which the interpretation and acceptance of the message can be negotiated by the recipient (Hall, 1980).

To sum up Jakobson’s speech event, with its six factors as the focus, is to realize the *integrativeness* of an act of communication. Bakhtin writes about the integrative features:

We know our native language – its lexical composition and grammatical structure – not from dictionaries and grammars but from concrete utterances that we hear and that we ourselves reproduce in live speech communication with people around us. We assimilate forms of language only in forms of utterances and in conjunction with these forms. The forms of language and the typical forms of utterances, that is, speech genres, enter our experience and our consciousness together, and in close connection with one another... (p. 83)

It is within the code and message, the selection and combination of words and phrases, with influence from the environment of the addresser and addressee and inclusion of the contact and context, that Jakobson’s model de-lineates speech into a necessarily dialogic act. For second language users and those without an intimate knowledge of the code and/or context, communication can become ‘tricky’ and the intended interpretation of the utterance can fail to be successful. However, if the message is presented in a multimodal

format, in which a second language user may choose the easiest, most familiar path to approximate comprehension of the message, then the success rate may improve.

1.7 The puzzle and comprehension...

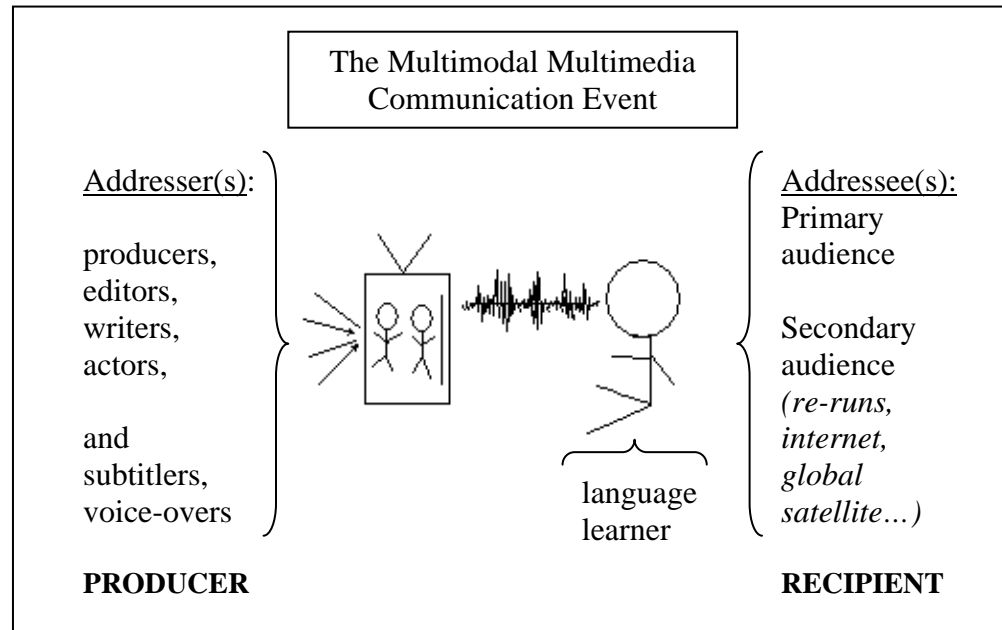
Taking the complexities of the functions of the speech event into consideration gives rise to a metaphor which I will attempt to provide and examine in the context of this study. If one takes the idea that comprehension of a message by the recipient entails interpretation of bits of the message as conveyed through different means, such as sound, gesture, color, the context, the speaker, the pretext of what happened before the utterance and the predictions of what might come next, and that these bits of information come in different modalities of presentation (sound, vision, touch, etc), then an interpreter of the message picks the bits that carry the most relevant and easily interpretable meaning *for that individual*. If a person is hard of hearing, then much of the message may come from the visual aspects of the message presentation, while if a person is a second language learner then perhaps the learner's strengths in, say, reading, may be the preferred manner of interpreting a message (therefore in a print medium). In other words, the meaning of the message is put together much like the pieces of the puzzle, whereby the interpreter of the message gathers all the information that he or she can and strategically puts them together to form the best, most complete picture, or interpretation, possible. Therefore, the metaphor that COMPREHENSION is a PUZZLE will be looked at from the point of a second language learner in a multimodal learning environment. Questions two and three will look at this metaphor in some detail using eye movements to look at the viewer's selection of information input, while Question three will use small case studies

to see if there is any correlation between a learner's strategies for language learning and what he or she actually does when watching a multimodal event.

1.8 Summary

This section has focused on setting up the ideas surrounding a Communication Event, drawing on Jakobson's Speech Event as the central schema. On the one side, Verschueren provides insight into the varied composition of the addresser(s), while on the other side exists the multiple layers of the addressee such as that posited by Goffman. However, the terms used by Jakobson are slightly different than those used by Goffman or by Verschueren, or by Kress & van Leeuwen or by other scholars who will be referenced in the forthcoming chapters. In order to truly place the addresser and addressee within the environment of a multimodal multimedia communication event, which is constructed of different producers of a message and multiple receivers of that message, a separate terminology will be used, as illustrated below in Figure 2.3:

Figure 1.3 The Multimodal Multimedia Event.



The Multimodal Multimedia Communication Event then is a media text that may be viewed by audiences other than those originally intended: a global community. Whether the original audience was an English speaking audience in Europe or North America, due to global communication viewers may exist in every time zone, years in the future, and could watch it in a variety of languages due to the addition of subtitles/closed captioning or voiced-over narration. Language learners (either L1 or L2) can use these media texts as tools to access the *code/langue* and the immediacy of the *message/parole* of the characters and story as it was originally intended for the primary audience, or the members of the original community. In addition, because of the multimodal aspects of the media text, learners have optional access points for interpreting and comprehending the actions and dialogue of the characters: gestures. Whether or not the gestures can be labeled as 'authentic', since they are performed by actors, the actions nevertheless are

Specker, Elizabeth. (2008). L1/L2 Eye movement reading of closed captioning: A multimodal analysis of multimodal use. ic.arc.losrios.edu/~speckee

acceptable to the primary audience and used by that audience as well to interpret and comprehend the media text. In that manner, they can also be used by a secondary audience to learn. In Chapter two, the Multimodal Multimedia Communication Event will be placed in relationship to this study and within relevant background literature.