



*The Reading Matrix: An International Online Journal*  
Volume 15, Number 2, September 2015

## **Creating Joint Attentional Frames and Pointing to Evidence in the Reading and Writing Process**

**John A. Unger**

Georgia Gwinnett College

**Rong Liu**

Georgia Gwinnett College

**Vicki A. Scullion**

University of Georgia

### **ABSTRACT**

*This theory-into-practice paper integrates Tomasello's concept of Joint Attentional Frames and well-known ideas related to the work of Russian psychologist, Lev Vygotsky, with more recent ideas from social semiotics. Classroom procedures for incorporating student-created Joint Attentional Frames into literacy lessons are explained by links to video data and a presentation of an exemplar case. This paper is intended to open several semi-structured, flexible avenues for systematically and collectively researching and integrating digital video cameras into classroom language/literacy learning practice and into the wider adult education, pre-college education, and ESL/EFL communities. The action-research case-study methodology presents an exemplar case who illustrates how learners use digital video cameras to organize and present Speech, a Visual, and the Act of Pointing to create and evaluate meaning, specifically abstract relationships of supporting details to main ideas during the process of writing a formal academic summary.*

### **INTRODUCTION**

Summarizing and responding to text is an important and ubiquitous communicative event that occurs across our personal, academic, and professional lives. We gather numbers, words, symbols, and all manner of digital images and sounds, and then organize and summarize this cacophony. This summarized information allows us to problem-solve and complete goal-oriented activities. This theory-into-practice paper reports on current progress and presents an exemplar case from a long-term, ongoing action-research oriented project with the overall goal of introducing digital video cameras into English language reading and writing instruction across a wide variety of educational contexts..

Through a case study (Yin, 2003) and action research design (Stringer, 2014), we are focusing on some of the immediate goals, which include examining the development of students' abilities to put ideas and chunks of language (i.e. approximately three- to nine-word segments;

formulaic or non-formulaic) together into meaningful summaries and responses. In other words, we are taking an approach to understanding learner interaction with language that emphasizes how students are intentionally and unintentionally positioning ideas and language in concrete and abstract ways to create meaning.

Research Questions Guiding the Study:

1. How do participants combine different concrete and abstract resources and ideas (e.g., their bodies, objects in the immediate environment, language, visuals, and directions for formal academic goals) to create oral and written summaries?
2. Is evidence of student development of explicit awareness of abstract ideas, such as the relationship of supporting details to main ideas and thesis statements, found in the video data?
3. Is there evidence of this development across modes (speech, the visual, chunks, reflections, and drafts of summaries) on the video-recorded presentations and in the final outcome?

Following some broad background information on digital video cameras and a presentation of the theoretical foundation, Case Larry is presented to illustrate the directions, concept development, and an exemplar outcome. When stakeholders (i.e. teachers, administrators, and researchers) understand the connection between the directions and the outcome, the better equipped they will be to make necessary adjustments for their own learners and educational contexts. (All video data and materials are available at: [http://transitional-literacy.org/?page\\_id=10002](http://transitional-literacy.org/?page_id=10002) password rabbit14)

## **BACKGROUND**

Under the broad array of computer-mediated communication, the use of digital video cameras for learning is similar to older computer-mediated activities with video options added as cameras have become more common; however, the major emphasis of computer-mediated language learning for teaching foreign languages seems to be the negotiation of meaning across modes offered by computers (Mahdi, 2014). Particularly for reading and writing, a wide range of options for using digital video cameras for instruction thrives at all levels of instruction.

The well-known activity of digital storytelling seems to be the largest area of digital video camera use that involves writing. Digital storytelling places an emphasis on positioning the student as teacher, known as reciprocal teaching (Gruenbaum, 2012; Palinscar and Brown, 1984; see also Wertsch, 1998). Learners are positioned to teach content as they are in the process of learning. A search of the internet and academic databases reveals a wide array of suggested methods of digital storytelling, which position the students as teaching, or at a minimum, revising and repackaging content in such a way as to learn the content through the process of collecting information, organizing information, and presenting information to an audience. Throughout the process, students work across modes and media to produce academic writing and content.

With regards to using a digital storytelling format in a more formal content area to prompt language learning across modes, Ranker (2008) reports on two 12-year-old students creating a video about the Dominican Republic from their own research. The two students in the Ranker study were struggling, and engaging them with language was a prime motivating factor behind the project. The data presented in the study clearly shows the two participants steadily working to read, write, and think in a multi-modal format (see also Kress, 2003; Kress & van Leeuwen, 2006).

As Ranker (*ibid*) and others have reported (Dumova, 2008; Hativah, 2000; Yoon 2012), the use of the digital story-telling format can motivate inquiry and language learning.

Gregory, Steelman, and Caverly (2009) reported on a number of options in developmental education contexts, such as personal narratives, outlining arguments, and doing other short 3 to 5 minute videos to integrate video into language learning. As with other suggested digital video strategies, students are responsible for producing edited video as one of the outcomes and work extensively with scripts, adding music, inserting images, and all the other tasks associated with video production. A different twist on digital storytelling was reported by Xu, Park, and Baek (2011), who had undergraduate students use the virtual world Second Life as a platform for creating digital stories. Using the story topic “Traveling through space and time,” students worked in teams to create a story. Students traveled through different worlds within the virtual world, and these travels become their stories. The students using the Second Life platform made gains in self-efficacy in writing and in a concept of flow, (see Csikszentmihalyi 1990), which emphasizes a level of comfort and smooth feeling when engaged in activity.

In another study positioning video into an active feedback loop, McNulty and Lazarevic (2012) found that students producing digital videos were highly motivated and had improved pronunciation and presentation skills at the end of the process. As with many of the digital video camera activities, a step was involved where students were prompted to evaluate their use of language and image to create meaning (see also Strassman & O’Connell, 2007).

One of the challenges of using video technology that is reflected in the current literature is the time commitment required to learn the editing software (see Ranker 2008). We have seen this in our own work with digital videos in the classroom. Students require additional time beyond what might normally be expected to produce a digital video in order to learn to use the technology, especially if they are involved in editing the videos using the kind of higher-level technical work (e.g. editing with video software) that is an aspect of other studies (Mills, 2010; Strassman & O’Connell, 2007). However, we have generally reduced the class time considerably (i.e., two one-hour and forty-five minute time periods at most; sometimes less) by avoiding the editing process altogether; we will return to this particular point later.

An essay from Reinders (2014) reported on different types of touch, gesture, and eye-movements involved in different software applications, as with touch screens. The way Reinders describes touch and gesture with digital devices has some relationship to our emphasis on the act of pointing at different phases in the activities. However, Reinders emphasizes more of the tactile and physical features of interaction than our research, which looks more at gesture as a part of overall semiotic design of meaning (see Kress, 2003). Reinders (*ibid*) cites older research on TPR (Total Physical Response) (Asher, 1977 is cited), where students respond to commands. Reinder reports that research into study classroom interactions is limited, but she makes a number suggestions related to segmenting text in different arrangements to highlight clause and phrase patterns; for the digital video activities in this study, dividing sentences and clauses into chunks of language is an essential part of arranging language on visuals (see Figure 3; see also Unger & Walter, 2010; Unger & Scullion, 2013).

The use of digital video to create documentaries and to present narratives seems to be the main option chosen for explicitly repackaging of information and working with language across modes when instructional goals are linked to the improvement of reading and writing. The literacy activities using digital video cameras found in the literature consist mainly of digital storytelling types of activities, all of which seem to require the learning and use of video editing software; consequently, large investments of time are involved. As far as we know, no current research

focuses on specific positionings of learner, speech, writing, text, and video cameras in a triadic arrangement to prompt the creation of idealized Joint Attentional Frames, which we will present shortly (Dumova, 2008; Fiorentino, 2004; Yerrick, Ross, & Molebash, 2005). However, all digital video activities, particularly where students have control over digital video cameras, have the potential to be understood and assessed by all stakeholders through the framework and time-efficient basic procedures presented in this paper.

## **THEORETICAL APPROACH**

The broad approach for language and signs used to guide and research the classroom activities presented in this paper is a synthesis of a number of well-known Vygotskian and semiotic theories (e.g. social semiotics; activity theory) (Davydov, 1999, van Leeuwen, 2005; Wertch, 1998; van Lier, 2004).

### **Joint Attentional Frames and Intention Reading**

Tomasello's (2003) concept of Joint Attentional Frames is based on the idea that humans have evolved species-specific abilities to read the intentions of others; language acquisition is prompted by these intention-reading abilities. A Joint Attentional Frame is created during social interaction that directs interlocutors' attention to some type of third entity (i.e. a triadic in which two or more interlocutors' attention is focused on concrete or abstract objects, contexts, and actions). Tomasello uses an example of a baby interacting with an adult to explain the concept. To paraphrase Tomasello's (2003) example, suppose I come into a room holding a bib. As I look at the bib and the baby follows my gaze to the bib, the baby understands that it is time to put the bib on and eat. Suppose I come into the room later with a baby stroller; the baby and I look at the stroller, and we both understand that the next sequence of events that will occur will involve the usual steps taken to get ready for an excursion out of the house. The bib and the stroller, third entities, are crucial features of intention reading.

In these examples with the baby, the interlocutors' focus on the bib or the stroller implies different meanings and actions as we create a Joint Attentional Frame. According to Tomasello, this is how language acquisition occurs; learning and conscious attention are based on our social interactions and our ability to create episodes of shared attention, and not on any universal grammar coded in our genes (see Tomasello, 2003: 21-23). Tomasello describes the importance and implications of this concept as follows: "The basic point is that joint attentional frames are defined intentionally, that is, they gain their identity and coherence from the child's and the adult's understandings of 'what we are doing' in terms of the goal-directed activities in which we are engaged" (Tomasello, 2003: 22).

This creation of Joint Attentional Frames supports language acquisition by prompting the child "to create a common ground within which she may understand the adult's communicative intentions when the adult uses a novel piece of language—at least partly by creating a domain of 'current relevance'" (p. 22). Of course, with adults, this situation evolves in a complex manner, although the basic principles involved of establishing shared meaning through a social act of pointing and shared intentionality are broadly the same.

To paraphrase another example from Tomasello (2003), suppose I go to a small bus station in a foreign country where few interact with English speakers, and I want to buy a ticket. If I ask

someone standing outside the bus station how much a ticket to a specific destination costs, they would probably not understand much. However, if I walked into the bus station and up to the ticket counter and was able to point to a destination on a sign and my wallet, or to an amount written down on a sheet of paper in the more common sign system of numbers, or point in the general direction of the destination, or to the clock on the wall, and follow this with other kinds of pointing (see deictics and pointing Kendon, 2004; Kita 2003) and mumblings and scribbles that go along with any interaction, I could make my intentions clear and successfully complete the buying of a ticket (Tomasello, *ibid*).

These general examples of interactions that create Joint Attentional Frames demonstrate the three interrelated concepts and observable activity that are emphasized in the ongoing research, namely Signification, Mediation, and the Act of Pointing.

## **Signification, Mediation and the Act of Pointing**

### **Signification**

According to Eco (1976) a sign is “everything that, on the grounds of a previously established social convention, can be taken as something standing for something else” (p.16 italics in original) (see also Peirce, 1991: 141). From this description, it is easy to see how the process of sign creation (i.e., signification) is easily one of the most ubiquitous acts in which humans engage. From our first baby moments of reaching for something and having this gesture collaboratively transformed into pointing (Vygotsky, 1978, p. 56; Tomasello, 2003), to later moments of learning algebraic formulas, complex sentence structures, and accompanying rhetorical strategies, this assigning of meaning to the world, objects, ourselves, and others creates complex representational systems humans use to complete goal-oriented activity in the world. This is well known, but what many do not realize is that one of the major characteristics that distinguish humans from other species is how we use representational systems to share intentions (Tomasello, 2003).

From the examples described of the baby and adult with the bib or stroller, or with the tourist in the bus station, or with an infinite number of events with signs that are a part of daily life, it is easy to see how objects, signs, and pointing are used to adjust and share our mental activity while mutually affecting the mental states of others and contexts in which goal-oriented activity takes place; human cognition, perception, and activity are mediated (Kozulin, 1998; Wertsch, 1998; Wertsch, 2007).

### **Mediation**

Humans mediate meaning through a continuous process of assigning meaning to object, events, and themselves in relation to others and the environment (i.e., the process of signification) (Wertsch, 1998; 2007; see also Wittek & Habib, 2013). For the tourist in the bus station, the clock mediated specific information about time, the numbers scribbled on a sheet of paper mediated the amount needed to pay. For the baby, the bib signified that a meal would take place; with a mutual gaze, the bib mediated intentions without any words at all. Other common examples of objects and signs mediating activity include keyboards on computers, measuring cups, PlayStation or Xbox Controllers, and a red light at an intersection.

The list of mediational means is endless, and as we introduce increasingly complex categories and associated sign systems into learning and teaching, we need some way to inventory

and track development of these ways of creating meaning. Fortunately, scholarship in understanding how to inventory our complex interactions of signs and communicative activity is offered by the dynamic and growing field of social semiotics, from which the concept of semiotic resources can be used to inventory and track individuals' and groups' creations and use of mediational means across modes and communicative contexts (Kress, 2003; Kress & van Leeuwen, 2006; van Leeuwen, 2005).

### **The Act of Pointing**

Crucial features of any communicative event are different types of deictic displays (i.e. pointing), particularly hand gestures (Kendon, 2004; McNeil 2005). Though there are always concrete and abstract features of any interaction that highlight one thing or another, gestures in general, with pointing in particular, are easily observable, though gestures are complex and often misunderstood mediational means (McNeill, 2012; Wertsch, 1998). From McNeil (1992; 2005; 2012) and Kendon (2004), we are adapting concepts and definitions, and taking a broader approach to the study of gesture; specifically, McNeil's (2012) precise definition of gesticulations, a type of gesture that McNeil distinguishes from Kendon (2004) and others in terms of the speaker's conscious intention of gesturing, is beyond the scope of this classroom-based instructional research..

Deictic gestures, that is, pointing gestures (Kita 2003), can be either abstract or concrete. Someone might say "I live here in Amsterdam" and point at the floor as two people sit and talk in a cafe in Amsterdam. One of the interlocutors then might say "I used to live in London" and point in some general direction away from the room (McNeill, 1992; 2005).

For the Joint Attentional Scenes described between the child and adult, recall that eye gaze was a crucial feature of setting up the triadic relationship, as it was in the bus station as the traveler pointed to the clock on the wall and a number on a piece of paper. One of the major issues in the ongoing research is how deictic and other gestures point specifically to intended meaning, and the function of pointing to highlight different chunks of language (see also Goodwin, 2003).

One effective way to understand how participants use the Act of Pointing as a resource for creating and expressing meaning, and to bring together the concepts of Signification and Mediation as analytical tools for investigating literacy processes and design, is to understand how concrete and abstract resources are purposely transformed into signs, into semiotic resources (Kress, 2003; van Leeuwen, 2005) .

### **Semiotic Resources**

During any communicative event, such as the traveler in the bus station, actions and objects have a "theoretical semiotic potential" and an "actual semiotic potential" (van Leeuwen, 2005, p 4). Actions and objects comprise the potential to be signified due to "past uses and their potential uses" (ibid), such as the baby knowing a bib signifies eating or a stroller signifies an outing. Participants authoring meaning in any communicative event, consciously aware or not of assigning meaning to actions and objects, are continuously transforming actions and objects into resources for communication, semiotic resources. Of course, as has been acknowledged by Bakhtin (1986), Pierce (1991) and now van Leeuwen (ibid) and many others (e.g. Kress 2003; Kress & van Leeuwen, 2006), creating and interpreting signs is a dialogic, dynamic, and -- now in the twenty-first century -- more potentially transformative event than ever before. For our ongoing research,

the theoretical framework aims to highlight how participants create or arrange semiotic resources into the triadic-like space described by the concept of Joint Attentional Frames.

To summarize the theoretical perspective: what humans signify involves a collaborative intention-reading process, which is shaped by some kind of pointing. Through this dynamic process of signification, semiotic resources are collaboratively created to mediate activity. Ultimately, signification and semiotic resources become inseparable from perception and cognition through the process of mediation.

### **METHOD: OBSERVING THE TRIADIC OF SPEECH, A VISUAL, AND THE ACT OF POINTING**

Through an overall action-based research design and case study methodology (Stringer, 2014; Yin, 2009) we are attempting to create the most authentic, real-world approach that we can; that is, we are working from Vygotsky's fundamental idea that learning and development, particularly concept formation, unfolds "right before one's eyes" (Vygotsky, 1978). The case selected for this paper, Case Larry, is one of approximately 50 Case samples of video, audio, and subsequent formal writing. We are revising the digital video camera activities as we are tracking the development of language across approximately six English for Academic Purposes courses with students from a wide range of English language and formal schooling backgrounds with students immigrating to the U.S. anywhere from early childhood to a week or so before the semester begins. Students from Pakistan, Columbia, Brazil, Vietnam, China, Korea, Jordan, Iran, Thailand, Afghanistan, Haiti, Bangladesh Somalia, Kenya, and India are some of the mix of the language groups and ethnicities that are present in these classes.

For Case Larry and prior cases (see Unger & Liu, 2013; Unger & Scullion, 2013), our intention is to identify specific reference areas in the data where all stakeholders can make judgments about the process of signification, mediation, and the way participants design and enact Joint Attentional Frames.

We have found that responding to text with a summary or brainstorming for an essay can be framed for analysis, teaching, and assessment by closely tracking three salient features of any communicative event: 1) speech; 2) a visual; and 3) the act of pointing. The speech in the examples presented in this paper and on the website are part of students' reading and writing response work. In all communicative events, the visual can be concrete, such as Power Point slides, film, simple posters, or something in the immediate environment that interlocutors point to, or visuals can be abstract, such as a visual created with the hands as interlocutors talk an image into being, perhaps something all may have shared, such as where they were at the time of a celebration or national crises, like 9/11. The visuals that co-occur in the classroom activities we have used is created by simple poster paper and colored pens.

For instructional and assessment purposes, during the actual communicative event, the major part of the presentation that the interlocutor has the most control over is the act of pointing, whether this pointing is generated by the hands, words (i.e., for example, such as, these, those, to be specific), music, color, lights, font, or any other means available to point one or more interlocutors' attention to a specific author intention.

## The General Directions for the Video Cameras, Visuals, and Case Larry Two Basic Guide Questions to Finding the Main Idea

We begin the activity by presenting students with a resource such as a text, short video, or even an image. Then we ask them to answer two basic, well-known guide questions that can be found in almost any fundamental reading and writing textbook:

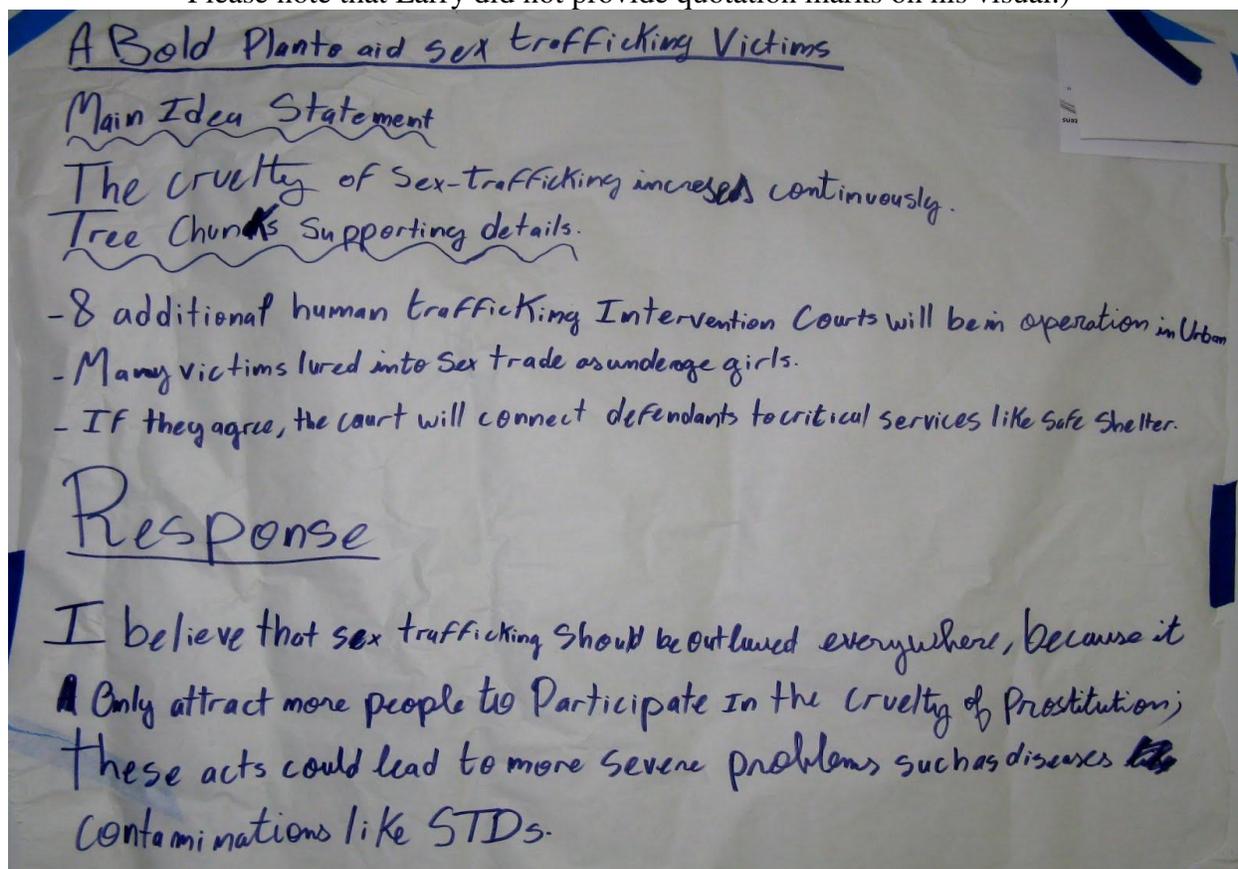
1. What is the topic?
2. What does the “author” want you to know?

With the second question, we make sure that students know that we can change the term “author” in that question to artist, director, musician, teacher, or any type of media or conversation in which an “intention” exists.

### Figure 1. The Visual Larry Made for Presentation

(It can be viewed at [http://transitional-literacy.org/?page\\_id=10002](http://transitional-literacy.org/?page_id=10002) password rabbit14.

Please note that Larry did not provide quotation marks on his visual.)



### The Steps Students Take to Create the Video Using their Visual and Pointing

Students are encouraged to follow the basic procedures below, which prompt the reading aloud of a supporting detail along with an explanation of how or why the supporting details are related to the main idea; a recent adjustment between the model video and Larry’s video is the addition of a response. Although not included in the written directions, students are frequently reminded

to point at chunks of language on their visual, but are also reminded not to over-do it. Also, pointing is identified on the rubric (see: [http://transitional-literacy.org/?attachment\\_id=9988](http://transitional-literacy.org/?attachment_id=9988) for the Excel Doc).

### **Directions:**

1. Read your original Main Idea Statement; then  
Introduce your supporting details in sequence by saying the phrases below that are in quotes; the capital letters below highlight the material on your poster paper that you should read:  
“The first supporting detail is”: READ THE SUPPORTING DETAIL  
“This supporting detail supports the main idea because”: SAY WHY YOU THINK THIS SUPPORTING DETAIL IS RELATED TO THE MAIN IDEA
2. “The second supporting detail is”: READ THE SECOND SUPPORTING DETAIL  
“This supporting detail supports the main idea because”: SAY WHY YOU THINK THIS SUPPORTING DETAIL IS RELATED TO THE MAIN IDEA
3. “The third supporting detail is”: READ THE THIRD SUPPORTING DETAIL  
This supporting detail supports the main idea because: SAY WHY YOU THINK THIS SUPPORTING DETAIL IS RELATED TO THE MAIN IDEA
4. Read your response statements; try to keep this only one or two statements (remember you want to write about a 4 to 7 sentence response on the final).
5. Choose the most appropriate supporting detail that you think supports your response: SAY WHY YOU THINK THIS SUPPORTING DETAIL IS RELATED TO YOUR RESPONSE
6. Conclude by saying anything you want, though if you are stuck with something today, say something like: “And that concludes my Main Idea and Supporting Detail Presentation about”  
SAY YOUR THEME HERE.
7. After the person concludes, the cameraperson should turn off the camera.

### **THE DISCOURSE DATA: CASE LARRY**

Recall the three broad research questions guiding this and other related studies:

1. How do participants combine different concrete and abstract resources and ideas (e.g., their bodies, objects in the immediate environment, language, visuals, and directions for formal academic goals) to create oral and written summaries?
2. With regards to participants’ development of explicit awareness of abstract ideas, such as the relationship of supporting details to main ideas and thesis statements, what evidence of awareness of the arrangement of information can be found in the video data?
3. Can we track this development across modes (speech, the visual, chunks, reflections, and drafts of summaries) in what we see unfolding on the video-recorded presentations and final outcome?

The data segment presented is Case Larry. Larry followed the current iteration of the procedures presented in this paper; of course, some variance occurs, and in most cases, is encouraged.

Larry chose an editorial about sex trafficking, followed the general guidelines for generating a main idea statement and writing chunks of text on his visual (see Figure One; The

Editorial can be found at: <http://www.nytimes.com/2013/09/26/opinion/a-bold-plan-to-aid-sex-trafficking-victims.html?ref=editorials&r=0>;

As per the directions, Larry started by answering the two guide questions, essentially misunderstanding some basic information from the reading. For this first step, as displayed in Figure Two and on the transitional-literacy.org page for this paper, Larry stated the topic as, “Sex-trafficking has increased in the last years,” and the author intention as “That prostitution is being practiced illegally without permit or approval from some states.” Although the topic could be understood broadly as what Larry proposed, the author intention (i.e. what does the author want you to know?) is completely misstated, as is the main idea statement: “The cruelty of Sex-Trafficking increases continuously.” When Larry moves to presenting his main idea statements and supporting details on video, he vividly exhibits difficulty in making his intentions clear. These tensions he displays are later noted by Larry and he makes adjustments.

For presenting the transcription in an accessible manner, the overall utterance was divided up into chunks that were differentiated by other chunks through a pause of some kind, or pointing. Description of the hand movements and other salient features of the interaction are written in italics below segments of transcription. Chunks that are important are underlined or placed in brackets (Also see the Video at [http://transitional-literacy.org/?page\\_id=10002](http://transitional-literacy.org/?page_id=10002) password rabbit14)

Transcript: Case Larry:

I will be doing na

A Bold na

Plan to end sex trafficking victims

Alright, my main idea statements is

the cruelty of sex trafficking increases continuously

and then my first supporting detail is

eight additional human trafficking intervention courts will be in operation in urban

And I believe this is ..supporting is related to the main idea

*As he is saying “this is . . . supporting detail,” he’s running his finger underneath the chunk “eight additional human trafficking,” back and forth, first to the end of human, then all the way to the end of trafficking*

Because there gonna be

There will be in an intervention

An operation

In Urban about about ah

sex trafficking

*As he goes through the four line sequence above, he points at the underlined words separately, isolating those words from the supporting detail statement.*

They’re kind of related. .

*As he says the above phrase, “They’re kind of related,” he waves his hand back and forth between the chunks and does three circular spins with his hand off to the side of the visual before saying “Yeah” and he scoffs, noticeably.*

Yeah

Alright many

And my second supporting detail is um

Many victims lured into sex trade as underage girls

And this

My second supporting detail is related to my main idea as well

Cause it talks about

*His index finger does a kind of half circle before landing on and underlining the word "sex trade" with his finger when he says "sex trading in the line below"*

Sex trading . . (points to the word "sex trade")

And victims (points to the word "victims")

Which is. .

A Cruelty (points to the word; he actually taps twice, close to the syllable emphasis) "cruelty"

And my third supporting detail is

If they agree the court will connect defendants to critical services, like Safe Shelter

And that one as well ah

Is related to my my main idea

Because it's a way we can solve this problem by

Taking it to court and doing all this stuff

*His hand is waving back and forth, like he is waving something away*

On the video, (see [http://transitional-literacy.org/?page\\_id=10002](http://transitional-literacy.org/?page_id=10002) password rabbit14) Larry can be seen and heard having difficulty expressing the relationships between his first supporting detail and the main idea. He can be seen bouncing his hand up and down, more or less waving away an explanation of the relationship between the main idea and the supporting detail; Larry is in the process of noticing this misalignment of information.

Pointing also dominated other moments in the data as Larry was putting together different chunks of text, some chunks as short as one word, to highlight specific themes, which evolved later and are seen in parts of his summary, such as when he puts together "sex trade" with "victims" and "cruelty." The very next line he reads becomes a solution to the tensions in alignments with supporting details and the main idea. Information from this line needs to be transformed from the supporting detail to the main idea position in the summary: "If they agree the court will connect defendants to critical services, like Safe Shelter."

Immediately after reading that third major point, Larry reveals the transformation in his thinking as he begins to express and develop the true main idea:

And that one as well ah

Is related to my my main idea

Because it's a way we can solve this problem by

Taking it to court and doing all this stuff

In the video, right before our eyes (see Vygotsky, 1978, p. 61), we can see Larry putting together different ideas, with tensions exhibited in the beats and his inability to finish the

explanation; at the end of the video, before he does his response, he resolves these tensions through stating a closer approximation of the main idea. Then in the later Self-Evaluation Step in the process in which students watch the video and answer questions, he reports his thinking on the error. (click on [Qualitative Response Framework](#) here or on the webpage and see Larry's response on the webpage). This is what Larry wrote:

My main Idea was clear but not very effective to my editorial because it give a bold emphasize of Sex-Trafficking, however, the editorial is about a “plan” to aid the victims of sex-trafficking, but my main idea sounds more like a major detail instead of main Idea, and If I would change it to something better I would write: A plan to help victims from sex-trafficking and stop the spread of illegal prostitutions.

Recall that Larry's original main idea statement was “the cruelty of sex trafficking increases continuously.” Larry is an exemplar case for how the process should lead students into explaining abstract relationships. He clearly exhibits difficulty in explaining the relationships of the main idea statement, which cannot work; this statement simply does not have a supportive relationship and a relationship cannot be created. Larry completely transformed his positioning of chunks and creates the following as a final outcome in the summary for the first draft he submitted: “Sex-Trafficking has increased in the recent years and the victims should rather be treated than being exploited with harsh court charges.”

## DISCUSSION

Case Larry demonstrates a continual process of signification and mediation that unfolds in any communicative event (see also Kress & van Leeuwen, 2006). Larry and other cases (see Unger and Scullion, 2013; Unger, Liu, & Scullion, 2015) present a transparent and ongoing development of speech, the visual, the act of pointing, and chunks of text; these units of analyses can be understood as semiotic resources. These semiotic resources gain and shed meaning depending on context, and through this process of pointing, different levels of intended meanings can be identified, thus opening up a window into specific learner process features for assessment as semiotic resources successfully create and express meaning, or not.

Larry's main idea statement was flawed, and as demonstrated in the data, his Joint Attentional Frame begins to break down because when he moved to share his intentions, and he realized that the supporting details are NOT related to the main idea statement, he begins to produce a cluster of up and down movements with his hands (i.e., *beats*, see McNeill, 2005), and at one point, waves away any more explanation. However, there is evidence that Larry resolves his tension with this flaw during the video process by transforming his thinking and creating a new main idea statement. As with some of the other studies presented about digital storytelling (Ranker, 2008; McNulty and Lazarevic 2012), the increase in the awareness of literacy strategies prompts a revision, and as with the digital storytelling in *Second Life*, these kinds of strategies demonstrate an increase in writing efficacy and flow; each example of using digital video had some positive influence in the classroom (see also Mahdi, 2014).

A common positioning of the student as teacher with reciprocal teaching emphasized is one of the main features of the current proposed use of the digital video camera with Case Larry

and others (see Unger & Liu, 2013; Unger & Scullion, 2013). Moreover, a focus on this specific positioning of resources during the fluid dynamic of communicative events might be a conscious way to further enhance the strength of Joint Attentional Frames as idealized metaphors to investigate speech, a visual, chunks of language, and the act of pointing, as these semiotic resources create meaning as a function of design (see Kress, 2003, for more on literacy as design).

### **IMPLICATIONS FOR CLASSROOM INSTRUCTION**

As we have been steadily collecting data and publishing different perspectives on different cases, several patterns have been found while integrating the digital cameras across classes and types of video (reading, summarizing, and presenting information). Overall, as mentioned earlier, the class time for the completing any of the digital video camera activities usually takes two one-hour and forty-five minutes sessions. We usually start with paragraphs where the main idea is fairly explicitly stated, then move to implicit-main idea paragraphs, and sometimes have students work with songs, for which the interpretations can be very challenging. The point is to give the students plenty of time to learn how to create information in limited camera frames and prompt them to play with language across modes with easier reading and writing assignments as they learn this approach and deal with the different ways they are prompted to think about language. Once students learn to use the cameras and the process, the reading and writing assignments can become more challenging; students do become efficient. I've had groups of 15 or 20 students create the visuals, complete the videos, and upload to the LMS system in one, one-hour and forty-five minute period. All the writing steps and a sketch of the arrangement of chunks of language are completed out of class and revised in class while putting the information on the Visuals. We are fortunate to use a Learning Management System (LMS) that allows students to upload and archive videos. At several institutions, due to LMS or internet limitations, flash drives were used along with an inexpensive ACER Netbook to temporarily store videos; the objective is for the student and instructor each to have access. Having the classroom space is also at times a constraint that has prompted us to have groups work in four corners of a room; a group in each corner. Students yell "quiet on the set" and they take turns; other groups either watch the presenter or work on their visuals. We are now taking the digital video cameras into a developmental math course, and are confident the entire process can be, and is intended to be, flexible across academic disciplines and applied technologies (see Unger, Liu, and Scullion, 2015).

### **LIMITATIONS AND AVENUES FOR FURTHER RESEARCH**

As with all research, there are limitations. One of the main limitations is our efforts to approach the analyses as teachers would in any classroom in the midst of a chaotic teaching day; consequently, we are sacrificing the fine analyses that could be undertaken with more time and technology. Moreover, this is case study research, so the findings have limited generalizability. However, because we are providing the reader with access to our data, readers and all stakeholders should be able to follow our path from data collection, to application and analyses, and to conclusions.

With regards to future research, we expect to eventually have well over 100 cases to archive and begin to closely examine statistical areas for assessment and begin to identify changing

patterns in literacy levels of students coming into and exiting the EAP program. Another area to research is to find out how learners feel personally about the videos and try to get them more involved in adjustments. Further research is also needed to create a dynamic assessment instrument to capture the literacy process unfolding in the video data, and to use this data in a more precise manner for placement into and exit from the EAP program.

## CONCLUSION

This theory-into-practice paper presents a synthesis and application of seminal Vygotskian concepts with more recent ideas from semiotics, language learning, and literacy development for the 21st century (Kress, 2003; van Lier, 2004; van Leeuwen, 2005; Wertsch, 1998; 2007; Vygotsky, 1978). By identifying the reference areas of Speech, the Visual, the Act of Pointing, and Chunks of Text in the data as these express Joint Attentional Frames, all stakeholders can grasp three important features of language and literacy development:

1. How meaning and the expressing of intentions becomes a function of design.
2. How participants' increased awareness of strategies and concepts of design positively influence the creation of meaning.
3. All stakeholders can track literacy development across different modes (e.g., speech, formal writing, and image) and gain a more authentic frame of reference in which to understand other measures of literacy to make decisions about avenues for classroom literacy instruction.

This critical literacy process of articulating propositions and evidence is a challenge for students in a wide array of educational contexts in the United States and abroad. Tomesello's (2003) ideas, the related Vygotskian theoretical perspectives, and recent work in social semiotics, along with digital video and access to related materials on the internet, offer new approaches to assist educators in increasing the effectiveness of critical literacy education.

**John A. Unger, Ph.D.** is an Associate Professor of English for Academic Purposes at Georgia Gwinnett College in Lawrenceville, Georgia, USA. His research interests include semiotic and ecological approaches to language and literacy instruction and how technology is used in the writing process.

Email: [junger@ymail.com](mailto:junger@ymail.com)

**Rong Liu, Ph.D.** is an Assistant Professor of English for Academic Purposes at Georgia Gwinnet College, in Lawrenceville, Georgia, USA. His research interests focus on language processing and using technology in second language reading and writing.

Email: [rliu@ggc.edu](mailto:rliu@ggc.edu)

**Vicki A. Scullion, Ed. S.** is a specialist in curriculum and instruction and a certified K-12 English teacher. Her current research centers on designing effective literacy lessons that incorporate technology for adult English for Academic Purposes students.

Email: [vascullion@gmail.com](mailto:vascullion@gmail.com)

## REFERENCES

- Asher, J. (1977). *Learning another language through actions: The complete guidebook*. Las Gatos, CA: Sky Oaks Production (2<sup>nd</sup> ed. 1982).
- Bakhtin, M. (1986). *Speech genres and other late essays*. Austin: University of Texas Press.
- Beach, R. (2012). Constructing digital learning commons in the literacy classroom. *Journal of Adolescent & Adult Literacy* 55 (5), 248-451.
- Csikszentmihalyi, M. (1990). *Flow: The Psychology of Optimal Experience*. New York: Haper & Row.
- Davydov, V. (1999). The content and unsolved problems of activity theory. In Y. Engestrom, R. Miettinen, and R. L. Punamaki (Eds.), *Perspectives on activity theory* (pp. 39-52). N.Y.: Cambridge University Press.
- Palinscar, A. S., & Brown, A. L. (1984). Recirpocal teaching of comprehension-fostering and comprehension monitoring activities. *Cognition and Instruction*, 1, 117-175.
- Eco, U. (1976). *A Theory of Semiotics*. Bloomington, IN: Indiana University Press
- Fiorentino, L. H. (2004). Digital video assignments: Focusing a new lens on teacher preparation programs. *Journal of Physical Education Recreation and Dance*, 75 (5), 47.
- Goodwin, C. (2003). Pointing as situated practice. In S. Kita (Ed.), *Pointing: Where language, culture, and cognition meet* (pp. 217-240). Mahwah, New Jersey: Lawrence Earlbaum Associates
- Gregory, G., Steelman, J., & Caverly, D. (2009). Techtalk: Digital Storytelling and Developmental Education. *Journal of Developmental Education*, 33 (2), 42-43.
- Gruenbaum, E.A. (2012). Common literacy struggles with college students: Using the reciprocal teaching technique. *Journal of College Reading and Learning*, 42 (2), 110-116.
- Hativah, N. (2000). *Teaching for effective learning in higher education*. Boston, MA: Kluwer Academic Publishers
- Kendon, A. (2004). *Gesture: Visible action as utterance*. Cambridge UK: Cambridge University Press.
- Kita, S. (2003). *Pointing: Where language, culture, and cognition meet*. Mahwah, New Jersey: Kress, G. (2003). *Literacy in the new media age*. New York, NY: Routledge Lawrence Earlbaum Associates Publishers.
- Kress, G. (2003). *Literacy in the new media age*. New York, NY: Routledge
- Kress, G., & van Leeuwen, T. (2006). *Reading images: The grammar of visual design* (2<sup>nd</sup> ed.).
- Kozulin, A. (1998). *Psychological tools: A sociocultural approach to education*. Cambridge MA: Harvard University Press.
- Mahdi, H.S. (2014). The impact of computer mediated communication environments on foreign language learning: A review of the literature. *Teaching English with Technology*, 14 (2) 67-86.  
Retrieved from: [http://www.tewtjournal.org/VOL\\_14/ISSE2/ARICLE4.pdf](http://www.tewtjournal.org/VOL_14/ISSE2/ARICLE4.pdf)
- McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago: The University of Chicago Press.
- McNeill, D. (2005). *Gesture & Thought*. Chicago: University of Chicago Press.
- McNeill, D. (2012). *How language began: Gesture and speech in human evolution*. Cambridge, U.K.: Cambridge University Press.
- McNulty, A., & Lazarevic (2012). Best practices in using video technology to promote second language acquisition. *Teaching English with Technology*, 12 (3), 49-61.  
Retrieved from: <http://www.tewtjournal.org/VOL%2012/ISSUE3/ARTICLE%204.pdf>

- Mills, K. (2010). Shrek Meets Vygotsky: Rethinking multimedia literacy practices in schools. *Journal of Adolescent and Adult Literacy*, 54 (1), 35-45.
- Peirce, C. S. (1991). *Peirce on Signs: Writings on Semiotic by Charles Sanders Peirce*, Ed James Hoopes. Chapel Hill: The University of North Carolina Press
- Ranker, J (2008). Making meaning on the screen: Digital video production about the Dominican Republic. *Journal of Adolescent and Adult Literacy*, 51 (5), 410-422.
- Reinders, H. (2014). Touch and gesture-based language learning: Some possible avenues for research and classroom practice. *Teaching English with Technology*, 14 (1) Retrieved from: <http://www.tewtjournal.org/VOL%2014/ISSUE%201/ARTICLE1.pdf>
- Smotrova T., & Lantolf, J. ( 2013). The function of gesture in lexically focused L2 instructional conversations. *The Modern Language Journal*, 97 (2) DOI: 10.1111/j.1540-1781.2013.12008.x0026-7902/13/397-416.
- Strassman S., & O'Connell, T. (2007). Authoring with video. *The Reading Teacher*, 61 (4) 330-333.
- Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.
- Unger, J. & Walter Gesture, speech, and graphic organizers as semiotic resources for summarizing: A two-case analysis of the genesis of meaning. *Asian EFL Journal*, 48. (13-54). Retrieved from: [http://www.asian-efl-journal.com/PTA/November\\_2010.pdf](http://www.asian-efl-journal.com/PTA/November_2010.pdf)
- Unger, J.A. & Liu, R. (2013). Digital Video Cameras, Main Ideas, and Supporting Details: The Process and Potential. *Journal of College Reading and Learning*, 44(1), 105-113.
- Unger, J. A., & Scullion, V. A. (2013). Digital video cameras for brainstorming and outlining: The process and potential. *Journal of Language and Literacy Education [Online]*, 9(2), 131-139. Retrieved from <http://jolle.coe.uga.edu/wp-content/uploads/2013/11/DigitalVideoCameras.pdf>
- Unger, J, Liu R, Scullion V. (2015). Language Competency as Semiotic Design: Attempting to Cross Academic Borders with Digital Video Cameras. *English Scholarship Beyond Borders*. 1(1), 1-51 Retrieved from <http://www.englishscholarsbeyondborders.org/wp-content/uploads/2014/12/Language-Competency-as-Semiotic-Design.pdf>
- van Lier, L. (2004). *The ecology and semiotics of language learning: A sociocultural perspective*. Boston: Kluwer Academic Publishers.
- Van Leeuwen, T. (2005). *Introducing Social Semiotics*. NY: Routledge
- Vygotsky, L.S. (1978). *Mind in Society*. Cambridge, MA: Harvard University Press.
- Yin, R. (2009). *Case study research: Design and methods* (4th ed.). Thousand Oaks, CA: Sage.
- Wartofsky, M. (1979). *Models: Representation and the scientific understanding*. Boston: D. Reidel Publishing Co
- Wertsch, J. (1998). *Mind as action* New York: Oxford University Press
- Wertsch, J (2007). Mediation. In H. Daniels, M. Cole, & J. Wertsch (Eds.), *The Cambridge companion to Vygotsky* (178-192). New York: Cambridge University Press.
- Yerrick, R., Ross, D., & Molebash, P. (2005). Too close for comfort: Real-time science teaching reflections via digital video editing. *Journal of Science Teacher Education*, 16 (4), 351-375.
- Yoon, T. (2012). Are you digitized? Ways to provide motivation for ELLs using digital storytelling. *International Journal of Research Studies in Educational Technology*, 2 (1), 25-34.

Xu, Y., Hyungsung, P., & Baek, Y. (2011). A new approach toward digital storytelling: An activity focused on writing self efficacy in a virtual learning environment. *Educational Technology & Society, 14* (4), 181-191.