An Introduction to Technological Support for Writing

Nick Greer
University of Arizona
Tucson, AZ USA
ngreer@email.arizona.edu

Jaime Teevan
Microsoft Research
Redmond, WA USA
teevan@microsoft.com

Shamsi Iqbal
Microsoft Research
Redmond, WA USA
shamsi@microsoft.com

Writing is a, if not the, core competency that transcends our diverse landscape of departments, programs, and curricula. Whether we are lawyers, doctors, artists, educators, scientists, humanists, public health professionals, policy experts, theologians, or just intellectually curious, we are all readers and we are all writers.

Samuel Moulton
HILT Director of Educational Research & Assessment

INTRODUCTION

This paper provides a brief introduction to opportunities that exist for providing better technological support for writing. Developing the appropriate technology that adequately supports the complex processes of writing is difficult because writing requires fundamental but varied skills such as reading, analysis, reasoning, and communication. For this reason, we believe that writing tools designed to support these skills can provide a valuable lens by which to understand and explore interesting problems related to information work. While a number of tools have been developed to support the processes of writing involving editing, organization and formatting, little exists to support the underlying cognitive processes of writing and research shows that management of these processes make the difference between simple and sophisticated writing. We believe that this represents a worthwhile opportunity for research. Here we give an overview of why writing is important and what is currently known about the writing process, and then discuss how this leads to new ways to thinking about how to build tools to support the writing process.

WHY WRITING IS IMPORTANT

We begin with an overview of why writing is important, and show that writing is a worthwhile problem to tackle. Writing is complex process that occurs in many contexts, but we have chosen to give special attention to business writing, where the writing is primarily used to communicate for productivity purposes and thus offers the opportunity for significant potential productivity gains.

General Writing

Writing encodes and transfers information, which is ephemeral and unstable in spoken form, by transcribing language into text, a form that is durable across space and time. However, text is not just a repository for information, nor is its production a mechanical process. Writing is an act of meaning-making. Through writing, we solidify concepts that were previously hazy, challenge and transform existing knowledge, and construct new models of reality [47].

It is therefore not a stretch to argue writing is one of the—if not *the*—most important human technologies. Writing is how we create and evolve major institutions, as in *The Constitution of the United States*, and how we manage a myriad of smaller undertakings, from building grocery lists so we have what we need for dinner to texting friends and family so we can stay in touch when we are apart. Furthermore, writing is not just a tool for communication, but also means of self-expression and an art, domains that are often considered subjective and messy, which is to say, non-technological.

Like any powerful tool, writing is difficult to learn, teach, and, more generally, *do*. When we write, our brains engage in a startling range of interdependent functions including verbal fluency, free association, episodic memory, divergent thinking, and goal setting, to name only a few [46]. Because writing is so integrative, it often overloads our working memory, and writers can become overwhelmed juggling the complexities of many concurrent processes. This can be particularly true if a writer has a deficiency in any area of the writing process, whether that is learned skills like spelling and grammar, cognitive functions like task switching, or the physical act of writing itself [41]. It is no wonder that The National Assessment of Educational Progress reports that only 27% of high school students write at a level deemed "proficient" or better [33]. By the end of college this number drops to 26.2% [10], with fewer than half of college students feeling like their writing improved by graduation [4].

The interdependence and integration that make writing so difficult are also what make it a worthy problem area for educators, technology providers, and researchers. Electronic writing tools, ranging from Microsoft Word to Scrivner to Emacs, are fundamental software for any computer, and the invention of word processing is considered one of the most important software innovations [14]. The prevalence of writing, the richness of the writing experience, and people's propensity to collaborate while writing, make the domain a fascinating area for study for computer science researchers interested in human-computer interaction, natural language processing, computer supported cooperative work, and more. Writing tools must draw from a range of research areas, and

insights derived from studying writing are likely to impact many areas of these domains beyond writing as well.

We believe that viewing writing as a technology, not in lieu of, but in addition to, an art, can produce healthy change in the way we write. Despite the fact that writing and technology are so clearly codependent—again, pen and paper are technologies too—the writing process has not fundamentally evolved to respond to significant changes in the technology we use to write. Scholars in composition studies have long tried to reform how writing is taught, but many teachers still cling to rigid, outmoded tools like the 5-paragraph essay, which has been in existence since the 1870s [39]. The technology to support writing has not evolved either to support the underlying cognitive processes that are at the core of good writing. While technologies like the typewriter and word processor have made the writing process more efficient and have been shown to improve the overall quality of writing [5], these are achievements in mechanics and information processing, not the web of critical thinking and goal-setting that define sophisticated writing.

How might researchers and technology developers catalyze this change? We believe that leveraging new technologies such as algorithmic-mediated workflows, natural language processing, machine learning, and crowdsourcing can demystify, assist, and eventually transform the writing process. We also think that because writing is such a fundamental and integrative activity, researching its processes and developing technologies that support these processes will result in discoveries that impact and influence interrelated fields, ranging from cognitive psychology, composition studies, machine learning, second language acquisition studies, creative writing, and communications.

Business Writing

One area of writing that is particularly interesting to explore is that of business writing. It is obvious that business cannot occur without communication, much of which occurs through writing. Like writing, businesses are complex and integrative, and it is essential that communication be clear, accurate, and consistent. Business writing must navigate specific rhetorical situations [Appendix A] that, while not unique to business, are heightened and nuanced by the exigencies and social codes of the business world, including:

- Businesses are explicitly hierarchical, and writers must understand how to write across different levels of power: Communicating among peers, soliciting clients, persuading superiors and subordinates.
- Businesses are fundamentally collaborative, which means the writing process often requires coordination.
- Business is conducted under the pervasive yet loosely-defined code of "professionalism," which necessitates careful control of tone.
- Business writing values accurate communication over perfectly formed, stylized prose, which can lead to different levels of "completeness."
- Business writing must be concise, clipped and efficient. It tends to foreground and highlight its main points in abstracts, executive summaries, and headers so readers can read horizontally to estimate value before reading vertically to extract that value.

Emerging changes to the way business is done complicate these situations. New technologies like smartphones, videoconferencing, and social media, and the work patterns they enable—telecommuting, outsourcing, etc.—mean that business can be conducted more efficiently along many measures—speed, efficiency, or quality—but these improvements come with hidden costs. Business is increasingly fragmented and remote, making writing, the technology that makes language durable across space and time, must do that much more to remain durable [3]. For example, an increasing number of communication channels (e.g., Twitter) provide new ways for companies to reach their customers. But the abundance of these channels means companies must invest in channel experts (e.g., a social media coordinator), or in unifying these disparate channels (e.g., developing house style guides). Without intelligent investments in writing, businesses risk being drowned by the very technologies that should be helping them.

While writing is an essential part of business, it is also undervalued. Employers report that oral and written communication skills are equally (if not more) important than technical skills, but hire disproportionately based on technical qualifications [20]. A study conducted in 2004 by the National Commission on Writing [34] reports that American companies believe two thirds or more of their employees are sufficient writers, but also that one third or fewer of their employees possess the writing skills they value. The same study estimates American companies spend up to $3.1 billion annually to compensate for employee writing deficiencies in the form of skills-upgrade courses and work that must be redone due to communication errors.

General opinion towards business writing is similarly paradoxical. Most view business writing as circumlocuitous to the point of being passive aggressive, and its profusion of jargon, buzzwords, and acronyms obfuscate more than they elucidate. At its best, business writing is laughable, deserving target of the satire of *Office Space* and *Dilbert*. At its worst, business writing is bureaucratic and inhuman, if not outright dystopian. However, the very things that inspire mockery or devaluation are the precise reasons why business writing should be improved, or at least, better understood.

WHAT IS KNOWN ABOUT WRITING

In order to provide better writing support, we must begin by understanding it. This section reviews research from the field of Composition Studies, which considers how writing functions—socially, cognitively, creatively, rhetorically—and how it is learned and taught. In particular, we focus on research that considers writing as networks of cognitive processes and which identify the pros and cons of specific patterns and pathways through these networks. We then

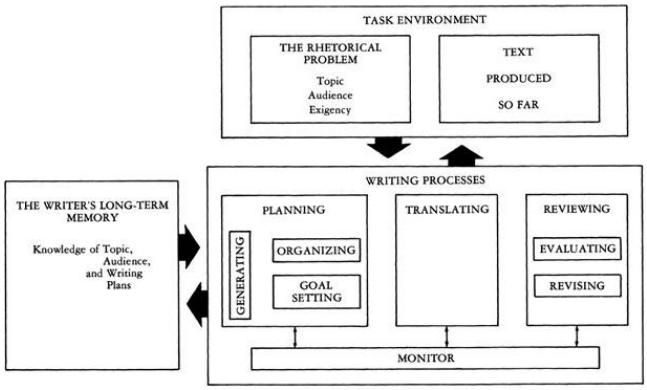

Figure 1. A model of Flowers and Hayes' Cognitive Process Theory of Writing [18], consisting of the Writer's Long Term Memory, the Task Environment, and the Writing Process.

recast these models in the context of Information Management to suggest specific writing strategies. We do this to further explore the complexities of writing networks and to anticipate our discussion of writing technologies.

The Writing Process

Writing is a complex act that resists being encapsulated in a single theory or model. Of the models that have been proposed, one of the more widely accepted ones considers writing as a network of cognitive processes, as shown in Figure 1. These processes are distinctive and hierarchical, but any process may be embedded in any other process, leading writers to construct highly personal, idiosyncratic, and, again, complex workflows [18]. The model consists of three aspects: the *Writer's Long Term Memory*, the *Task Environment*, and the *Writing Processes*. While each of these are essential to the act of writing, we choose to highlight the components of the Writing Processes because the processes of this area are the most explicit and distinct, and therefore the most readily modeled by a technology. (For more on the Rhetorical Situation, an important component of the Task Environment, see Appendix A.)

Writing Processes comprises four sub-processes: *planning*, *translating*, *reviewing*, and *monitoring*.

During **planning** writers form internal representations of the knowledge they will use in writing. Writers *generate* ideas by retrieving them from their long-term memory, then *organize* this knowledge by grouping ideas to create new ones or by identifying ideas that are weak and require additional retrieval phases. Deciding whether to group or perform additional retrievals is a product of *goal-setting*. Each aspect of planning can be about the content, or recursively about the writing processes themselves, and they can occur at any phase in the text's evolution, from the earliest stages when ideas are raw or hazy to later stages, such as ordering fully-formed paragraphs. Framed in the language of computing technology, the planning process resembles the growing architecture of a structured database. An architect identifies important groupings of raw data (tables) and links between groupings (keys). As these grow, the knowledge base stabilizes, allowing less taxing, higher-order retrievals (queries).

Because the representations built in the planning phase are abstractions, often taking the form of an image or a feeling, writers must **translate** them into text. Even representations that are word-based are more likely to be organized in loose networks rather than in linear, sequential prose with well-ordered syntax. Translation is both mental and physical; writers make word choices and orderings in their minds and inscribe these choices in the medium they are writing in. Writers who are deficient at any part of the translation process often find that that part overwhelms their short-term memory and precludes the more global planning processes from occurring. In an attempt to lessen this burden, many existing technologies solve for specific translation tasks (e.g., autocorrect solves for common spelling errors).

As ideas are organized and translated, they must be tested against the goals of the evolving text. This process, **reviewing,** often interrupts planning and translation, giving the writer a chance to *evaluate* their decisions and *revise* them if they do not fit their goals. This process applies to both prose and its pre-translated representations. For example, in diction, writers might run through a number of synonyms before actually writing down their choice. Many existing technologies also solve for specific reviewing tasks. For example, spellcheck evaluates strings against dictionary entries, flagging any strings that do not appear in the lookup, which prompts a writer to revise their spelling.

The last process, the **monitor**, is the meta-process by which a writer determines when to switch tasks and what task to switch to. These determinations are based on the goals of the specific text as well as the writer's habits and personal style. In the same way that reviewing can interrupt planning and translation, monitoring constantly interrupts other processes. There are many different processes to monitor, any of which can call another as a subroutine, so the monitor is constantly triggering across different functional areas and hierarchies. For example, choosing the word "big" instead of "large" is a translation task that might prompt a review of all other instances where this decision was made. This broader review process might then trigger a question about the text's goals against the reader's style. Technical writing values uniform diction, but writers often vary their diction for readability. Even when a global decision is made, it must be tested against local instances.

Writing Strategies

Writing strategies, or semi-rigid workflows that reduce the frantic monitoring process, are a necessary part of the writing paradigm. A writing strategy *decomposes* the complex task of writing and transforms it to smaller, more manageable subtasks (also called *microtasks*). Here we discuss some of the challenges with simplifying the writing process into a specific strategy, and then describe some writing strategies derived from information management.

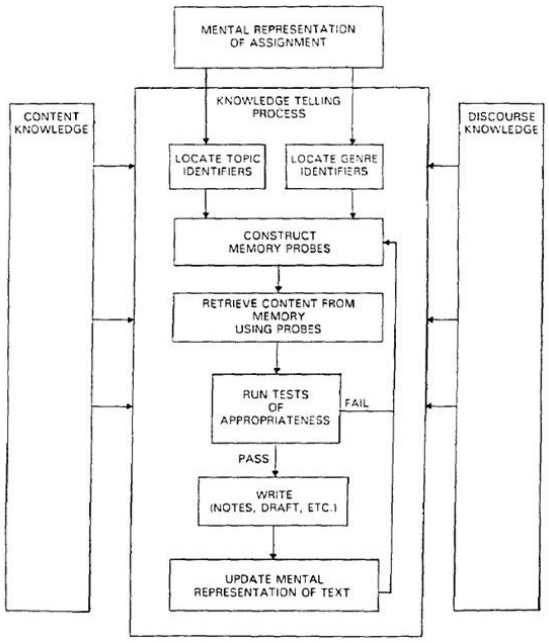

Figure 2. Knowledge-telling strategy [47].

Simplifying the Writing Process

Writing strategies enable writers to manage an otherwise unmanageable process, but risk being so rigid or simplistic that they lead to writing that misreads its rhetorical situation or fails to achieve its goals. An example that demonstrates this is the difference between a knowledge-telling strategy (Figure 2) and a knowledge-transforming strategy (Figure 3). Knowledge-telling is linear and sequential, resulting in writing that repeats what a writer already knows. Knowledge-transforming is highly recursive, integrating the new knowledge acquired in the course of the knowledge-telling process (which it calls as a subroutine) into its long-term memory and task environment.

The knowledge-telling strategy is fine if the writer's goal is simply to convey existing knowledge, but because writing is a space in which we make meaning, this is rarely the case. The knowledge-telling model insufficient for complex types of communication with complex goals. For example, a company reporting diminished earnings to its shareholders may seem like an instance of knowledge-telling, a simple conveyance of data, but because this data will impact the way shareholders perceive the company and the shareholders have some power over the company's success, suddenly the writing must consider and manage what might be unpredictable or irrational sensitivities. A knowledge-transforming strategy is much better equipped to handle this more complex rhetorical goal because it can continually compare problems of content (the accuracy of the data) against a rhetorical problem (the values of the audience).

We show these two strategies not because any one step is necessary to modeling the process of writing, but to illustrate the fact that there are different models that suit different rhetorical situations and writing styles, and while these

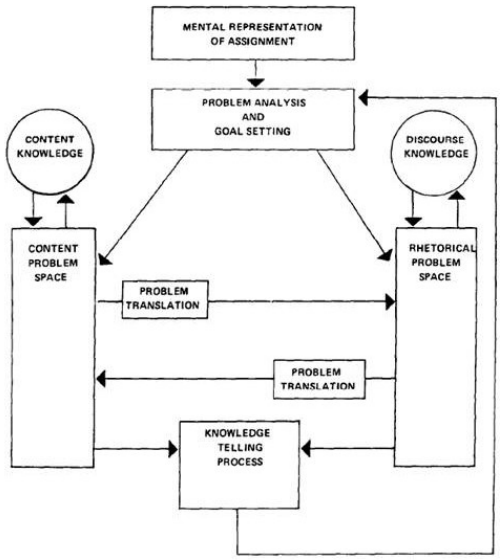

Figure 3. Knowledge-transforming strategy [47].

models are codifications and simplifications of the infinite graph of possibilities, each model is still extremely complex.

Writing Strategies from Information Management

While it is useful to think of writing as a series of physical and cognitive processes including issues of short-term memory capacity [40], productivity, collaboration, and multi-tasking, writing can also be viewed more generally as a form of *information management*, a discipline that provides a different and necessary lens into writing.

Five information management strategies have been shown to reduce cognitive load and improve writing quality [19, 29]:

- Chunking – breaking the process into manageable pieces with key points,
- Scaffolding – creating an organizing logic or narrative to the process,
- Pacing – slowing the process down to a rate that works for each writer,
- Interacting – communication between the writer and other judges (e.g., editor, example writing), and
- Monitoring – observing when writers have issues and intervening appropriately [30].

The word "process" is purposefully ambiguous to show how each can apply to any writing process or hierarchy of processes. For example, scaffolding can involve working on individual sentences before arranging them into complex paragraphs (i.e., organizing) or proofreading for typos and other technical flaws before considering a text's larger structure and flow (i.e., evaluating). Scaffolding can also involve checking that a single point within a paragraph coheres with the main point of that paragraph, or up a layer to the main point of a series of paragraphs. Scaffolding can even be a meta-process, as in the decision to do scaffolding before, after, or in addition to the other methods, to "create an organizing logic or narrative to the process."

The word "process" is also purposefully ambiguous to show that, while these strategies have been shown to improve writing, they are general ones that can improve learning and performance in other disciplines. When tailoring them to the unique aspects of writing—including language, text, inscription, social contexts, etc.—these strategies become especially powerful but also complicated. In addition to altering the physical and cognitive processes of writing, they also impact other interrelated dimensions. In the earlier scaffolding example, moving from less complex processes to more complex processes is not an abstract move, but is situated in with the text itself. As such, information about the text is also scaffolded, in this case, the length of a textual unit: words being (generally) less complex than phrases, which are less complex than sentences, and so on.

Implementing a Writing Strategy

The way each strategy gets implemented in practice can be impacted by many factors. Four important factors, precipitated from the three areas of the writing process model discussed earlier, are:

- Rhetorical – the rhetorical categories of writing (mechanics, organization, semantics)
- Stage – a text's development (pre-writing, writing, re-writing)
- Scale – different lengths of text (word, phrase, sentence, paragraph, section, etc.)
- Process – the physical and mental processes of writing (monitor, assess, transform, integrate)

Managing information with an eye for **rhetorical** categories has been shown to help the writing process, especially with weaker writers who spend the majority of their time and effort attending to surface-level *mechanical* details, rather than the more complex processes of meaning-making (*organization* and *semantics*) [42]. This distinction between mechanics and the other two areas influences the task's scale, with mechanical tasks typically concerning the word, phrase, and sentence, and organizational and semantic tasks occurring across multiple sentences and paragraphs.

Managing information by the text's developmental **stages** (*pre-writing*, *writing*, and *re-writing*) is valuable because it scaffolds the text's growth to mimic the three sub-processes we saw earlier with the writing process (planning, translation, and reviewing), prompting thoughtful engagement with tasks people would normally skip, and easing cognitive load [23]. However, these stages promote a writing process that is more linear and sequential than the hierarchical and recursive process good writing requires [46]. For example, if generating new ideas is sequestered to pre-writing, writers will not be able to integrate what they learn in the acts of writing and re-writing.

Less research has been done into the impact of information management on the other two factors, **scale** and **process**, but we speculate that all of these factors follow the Goldilocks principle, which dictates that conditions outside antipodal margins will diverge ("too hot," "too cold"), and conditions within will converge ("just right") [24]. For example, with scale, only considering a text's individual *words* makes a writer myopic to large scale units, but only considering whole *sections* makes it impossible to narrow in on specific *phrases* or *sentences*. Each task has margins within which it is effective and outside which it is not.

Additionally, each factor is interrelated. For example, conducting a final spellcheck is at once a mechanical task (rhetorical), a re-writing task (stage), a word-level task (scale), and an evaluation task (process). While most research considers individual factors to isolate its function and effects, we speculate the Goldilocks principle applies here too. Fixating on one factor, say, scale, by evaluating every letter, then every word, and so on, until evaluating the document is inefficient and masks rhetorical considerations like consonance and assonance (i.e., how the sounds of individual words repeat across phrases and sentences). The opposing extreme is to consider all factors and in doing so over-determine their codependence. For example, it is absurd to apply spellcheck to every stage of a text's development; in pre-writing there are no words to spellcheck.

Recomposing the Output of a Writing Strategy

Thus far we have discussed different strategies for breaking down the task of writing. However, decomposition is only half of the picture. Writers must also *recompose* the output of each subtask into a unified whole. Researchers have identified the following six recomposition strategies as necessary to knowledge-transforming [47]:

- Search – identify beliefs about a rhetorical situation,
- Delimit – restrict attention to a subset of ideas,
- Cohere – fit delimited ideas among one another,
- Fit – fit delimited ideas to a central point,
- Structure – build a schema for the ideas, and
- Review – review ideas across levels of hierarchy.

These strategies are complementary to the five information management strategies that decompose writing.

Applying the Right Writing Strategy

Though recomposition can mirror decomposition to some extent, neither is typically completely sequential, at least not with sophisticated writing. Recalling the knowledge-telling strategy, writers learn from the act of writing and will update their goals to reflect what they learn. This means that i) there is no perfect decomposition, and ii) writers must continually update their decomposition. For example, if a writer is writing on a timeline (pacing), but is slogging through a particularly difficult paragraph, it does not make sense to quit that paragraph after the time limit has been reached. The writer must either break the timeline and work through the paragraph (pacing) or identify this paragraph as exceptional (monitoring), flag it for later (scaffolding), and move onto the next paragraph (pacing). The original decomposition—pacing at the paragraph level—was too rigid for the rhetorical situation, so it prompted a sub-process.

This may seem a little circular or redundant, but that is because of the deep recursion of the writing process. Processes always have sub-processes. Tasks can always be further decomposed into microtasks. This graph is infinite, but our brains are finite, and the need to move from most general to most granular—from the writing process as a whole, to strategies that structure this process, and finally to decompositions and recompositions, the building blocks of good strategies—is a kind of decomposition in search of the scale of complexity where our brains say "just right."

TECHNOLOGY TO SUPPORT WRITING

In the context of what is known about the writing process, we now review computer technologies that support writing. We focus on new technologies that can enable knowledge-transforming via decomposition and recomposition, and consider who might perform each task when decomposed.

Current Technologies

While there are many computer technologies dedicated to writing, most of these address processes in isolation, not the web of critical thinking and goal-setting that defines sophisticated writing. For example, a tool like spellcheck is valuable because it can identify errors (evaluating), but only when paired with autocorrect is it able to fix these errors (revising). While spellcheck is undeniably more accurate at identifying spelling errors than the human eye, this is only for words in its dictionary (Long-Term Memory) meaning slang, jargon, and other neologisms that might be relevant to the text (Rhetorical Situation) could register as false positives. Even word processors, complex programs that manage many writing processes and have been shown to improve the overall quality of writing [5], achieve this affect by leveraging the comparative strengths of computing, not by emulating what humans do. Computers can read and write structured data at superhuman speed and volume, but fail when the structure changes or is too simplistic.

New writing technologies are starting to accommodate the variability and complexity of the writing process. To return to the spellcheck example, pairing spellcheck with predictive and corrective text expands it from purely evaluative to a tool that impacts planning and translation too. Spellcheck that integrates contextual data or machine learning can expand its dictionary to include non-dictionary words that make sense for the rhetorical situation and the text's history. In both cases, spellcheck improves by inflecting its isolated computational processes with aspects of the larger writing process and the human intelligence behind it.

In some cases this human intelligence is worked explicitly into the computational design, as with writing technologies that outsource microtasks to the crowd. When the crowd's (generally undifferentiated) labor is shaped by algorithmic-mediated workflows, the crowd can handle a range of tasks from small, mechanical tasks like shortening and proofreading [7] to larger, seemingly monolithic tasks like writing a news article [2], research paper [36] or novel [25]. We believe that technologies like these, which merge computation and human intelligence, have the potential to assist and ultimately transform the writing process. Technologies built on decomposition and recomposition, and the related concept of sourcing decomposed tasks, have the ability to mimic, augment, and replace not just isolated processes, but entire networks of processes.

Decomposition and Recomposition

Writing strategies take the large, complex task of writing and break it down into smaller tasks that can be completed individually and then recomposed. Because computers make frequent use of hierarchies and recursion, the strategies discussed earlier in the context of Composition Studies and Information Management extend readily into computational systems. While classroom decomposition is managed by an instructor and structured by lesson plans, and recomposition is executed by students, computers can support decomposition and recomposition algorithmically [50].

Research into algorithmic task decomposition looks at breaking large tasks down into subtasks that require minimal attention, context, and skill. This has been done for information tasks such as arithmetic [11], ontology creation [12], and caption-writing [27], where decomposition has been shown to support mobile work [35, 36], improve the quality of work, worker experience, and a worker's ability to bounce back from interruptions, at the cost of time [11]. The lattermost benefit, the bounce-back factor, is especially useful for business tasks, which are often conducted in small, discontinuous bursts or in collaboration with others.

Most successful breakdowns have been done for information tasks decompose into microtasks that require little to no context about other microtasks or the parent task. Writing, being a highly networked process, is context-heavy, which presents additional challenges to known complications of decomposition. Some writing tasks, such a spellcheck, are easy to decompose, but tasks that require special context, (e.g., write in house style) or that spawn a tree of subtasks (e.g., ordering paragraphs) are not easily decomposed.

How do we know which tasks require special treatment and which do not? Clues can be found in how writers already implicitly engage in the process of decomposition when they elect to, say, proofread rather than compose a new paragraph. This is the function of the monitor, the executive function of the writing process. Observing this function can produce models for when and how a task should be decomposed. For example, many writers prefer to scaffold by task type as a way of scaffolding for difficulty. First they perform simple mechanical tasks, such as spellchecking, and work their way up to complex semantic tasks, such as composing a thesis [9]. While these observations currently come from slow, manual data collection techniques such as verbal reports and protocol analysis [21], natural language processing and machine learning can be used to monitor a writer's strategies and habits on the fly as they and the text evolve [54].

The scaffolding example has a secondary benefit: it can help writers acquire context. Simpler, mechanical microtasks

require little context, but they can function as context for a more networked microtask. This is especially attractive because it communicates implicitly and kills two birds with one stone: the completion of one microtask and the decomposition of a context-related microtask from another parent task. The sequence in which writers handle microtasks has a bearing on efficacy and satisfaction [9], but in some cases it may be hard to replace the context of a parent task. For example, when evaluating an abstract concept on a large scale, such as the flow of an argument in a multi-page essay, the task can only be completed imagining the entire text.

Mimicking the tried and true decompositions of traditional writing might not be the most effective way to decompose writing when mediated by technology. A traditional writing decomposition might be a local optimum that masks a more global optimum. A successful tool will not simply identify the path of least resistance, but also new ways in which a task can be decomposed and recomposed even if its benefits are not obvious to the writer. A common example from Composition Studies is to add an explicit pre-writing phase (pacing the process) in which writers brainstorm and make outlines (planning), which has been shown to help writers who have a hard time knowing what to say or who dive into writing without planning ahead [23]. The MicroWriter, a program that decomposes the writing process into three types of tasks—idea generation, idea labeling, and writing—elects to further decompose one of its pre-writing tasks, idea labeling, into four microtasks—generate, merge, finalize, group—which allows more time for raw ideas to solidify and permits the automatic creation of outlines [49].

Collaborative Writing and Sourcing

The concepts of decomposition and recomposition are not limited to the writing process, or the *how* of writing, but can also be applied to the *who*: the writer. Writing, especially business writing, is often explicitly collaborative, with many writers involved with a single text. Writing is also implicitly collaborative. Even when there is a single name on the byline, a writer is never a single person. Our writing is always co-authored by the texts we have read or the instruction we have received, as well as the technologies we use to write, whether that is pen and paper or a word processor. We are constantly outsourcing.

Collaborative writing is a no less complex a process than traditional writing. With many discreet activities [45] and evolving roles [37], collaborative writing requires technologies that manage text transformations and meta-commentary in robust and novel ways. These needs are either handled asynchronously through the change tracking and version control features available in modern word processors, or managed synchronously through online collaborative authoring tools [38]. These tools explicitly facilitate coordination across authors by supporting awareness of what others are doing. For example, Tam and Greenberg [48] discuss the importance of workspace awareness drawing attention to certain changes via visual cues. Birnholtz et al. [8] argue that minimizing the visibility of some changes can facilitate social interaction. The MicroWriter takes this to an extreme and provides no direct awareness of other's writing actions within the tool [49].

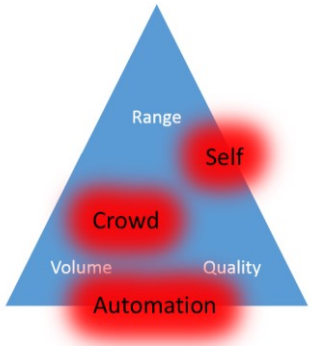

Figure 4. An illustration of the relative strengths of different labor sources for writing.

When tasks are decomposed to manage collaboration, we talk of them as being *sourced* to different labor pools. Tasks that create single-authored documents are sourced to the self [50], but they may also be sourced to the crowd [7], or even automated processes [52], each which has different strengths and weaknesses. For example, consider the general trade-offs to be made between the volume of tasks a source can handle, the quality of a source's work, and the range of task types that source can handle (Figure 4). The self has access to context and can handle a wide range of tasks well, but is subject to cognitive blocks. The crowd has limited access to context and can handle a narrower range of tasks, but can do so at scale. Automation is even better at handling volume, but provides the most narrow and rigid task support and is error-prone. However, these are only general tendencies of each labor source, and each merits deeper consideration.

Self

The self is the default labor source because writing tasks are rarely decomposed and separated from the writer. However, looking at the self through this new lens reveals a unique set of skills and limitations. The self is good at handling a wide range of tasks, but is error-prone and cannot handle large volumes of work with consistent results. What separates the self from other human agents is that the self typically has rich access to both domain and document-level context that others do not. However, this does not mean that the self is always the best at handling all of their tasks. Even strong writers are subject to the "curse of knowledge" [43] or "writer-based prose," [17] in which deep familiarity with a context or domain blinds the writer from seeing the text from a reader's perspective. Writer-based prose can be dense with jargon and makes leaps that appear illogical to the reader. Writers of all ability levels benefit from editors, which suggests the self could be augmented by other labor sources.

The self's biggest limitations are in motivation and work habits. Being a complex process, writing can be a source of frustration, even humiliation, especially in the performance-driven environments like school and work. Writers with low

or indeterminate self-efficacy tend to avoid the writing process [52], which often creates a negative feedback loop in which writers do not afford themselves enough time to write effectively. When presented with multiple tasks types, procrastinators choose short-term but lesser payoffs over higher, long-term payoffs [16]. Workflows that create interdependent short-term goals that build to higher, long-term payoffs (i.e., scaffolding) have been shown to help procrastinators by creating a "ramp up" process [1]. Scaffolding has been shown to help writers of all ability levels improve the quality of their writing.

Crowd

Traditional writing sources labor to other people whenever writers share their work with a reader, editor, or collaborator. Decomposition and sourcing facilitate these existing processes by reducing the coordination costs involved with sharing tasks subcomponents with collaborators [49], and, at a larger scale, paid crowd workers. Mechanical Turk has between 1,000 and 10,000 workers available at any given time [22] that can be used to support writing. Timeliness of response varies with the price point and the saturation of competing microtasks in the marketplace, but retainers, in which crowd employees are paid a bonus to wait and perform a future microtask quickly, can enable task completion in just seconds [6]. Crowd labor can be further coordinated into synchronous, collaborative groups, which improve both speed and quality, as in Chorus, which allows crowd workers to live-vote on tasks based on a shared working memory [28].

A number of algorithmic-mediated workflows have been explored to support the integration of crowdsourcing into the writing process [2, 7, 25, 36]. An important aspect of these workflows is that they allow for diagnostic and verification rounds. For example, Soylent's Find-Fix-Verify workflow asks workers to "verify" redundant "fix" tasks in an effort to identify higher quality work [7]. In traditional writing, quality is mostly a product of the skill and motivation of the writer, but algorithmic workflows make quality an output with many inputs. Still, the composition of the labor pool is essential. Some crowd labor platforms, such as Upwork, market their crowd workers as experts in specific categories and skills (e.g., copy editing). Others, such a Mechanical Turk, market workers with less differentiated skill sets but greater flexibility in the types of tasks they can support. Researchers who built Ensemble, a crowd creative writing platform, found that establishing clearly-defined jobs and power relations between roles improves motivation and leads to more successful final products [25]. Task type also has a bearing on speed and quality [53], with tasks like shortening individual sentences are more easily crowdsourced than ordering paragraphs in a multi-page document.

For some complex writing tasks, the context necessary for completing the task successfully must be communicated to the worker. This communication typically takes two forms: writer-to-crowd and crowd-to-crowd. Packaging tasks with instructions, rubrics, and examples (writer-to-crowd) are effective and efficient for communicating, but are unidirectional and do not always capture the required context [13]. Crowd-to-crowd communication (e.g., observing the behavior of other workers or passing information to other workers) has been shown to increase quality by transferring expertise and domain knowledge, though the quality of crowd-generated feedback is variable [13]. Context may also be communicated implicitly by repeating tasks in the same domain, document, or type of task [9].

Automation

The decomposition of a writing task into small pieces also makes it easy to incorporate automation. The microtasks that make up many algorithmic writing workflows are typically designed to require limited outside knowledge and contain as much of the required outside context as possible. As a result, these microtasks look much more like the kind of task that can be automated using machine learning than many typical writing tasks. For example, the MicroWriter [49] transforms the complex task of organizing written ideas into a series of microtasks that involve merely labeling a single idea or merging related labels. While state-of-the-art automated systems cannot organize text, they could accurately learn which label to apply to a piece of text which means that it is possible to automate much of the MicroWriter's organizational process. Successful natural-language processing techniques already exist to do a number of simple writing subtasks (e.g., identifying the tone of a sentence) that can be combined to complete more complex tasks (e.g., improving the tone of the entire document) [32].

It is likely only a subset of the microtasks associated with a complex writing task should be sourced to automated systems. One nice aspect of this is that it makes it feasible to consider hybrid workflows that incorporate imperfect automation into the writing processes by actively engaging humans to approve or reject the output of an automated process. Currently writing tools only incorporate automation when the outcome is very certain. For example, Word may automatically correct spelling or grammar errors, but it only does so when it is clear there has been an error and the system knows exactly how to fix it. Word does not do other types of corrections (such as changing a snarky sentence to improve the tone) because it is unlikely to do so well. However, by incorporating active human supervision it may be possible to extend the range of tasks that can be automated. A microtask, for example, could suggest several different versions of the same sentence automatically re-written in different tones and ask the user to select their favorite [26].

Additionally, as users perform each microtask the system collects valuable training data that could be used to eventually automate it, particularly for microtasks that are common across many different macro-tasks. For example, a microtask system for copyediting text could learn to identify sentences that need correction automatically by watching the author perform the identification for the first few pages. Using the data collected about how people complete these

microtasks, it is possible to learn and improve automated algorithms so to automate increasingly larger aspects of a task. When the data is collected across people, these models are generalizable; personalized models can also be created using the data collected from a single individual.

IMPLICATIONS

Writing is a technology made possible by other technologies. Writing allows language to be durable across space and time, but this durability varies with that of its supporting technologies. Think of a pen running out of ink or a word processor freezing, or a slower failure, such as the extinction of a language. And much like the processes of writing, the technologies of writing are highly networked and recursive. Any change to the technology will change the writing process, and vice versa. For example, smartphone adoption has led to developments in predictive and corrective typing to compensate for frequent errors on the small, mutable keyboard that provides minimal haptic feedback. However, as corrective typing becomes a norm, so too will the errors corrected by such a system, further entrenching the need for what was originally a secondary technology.

When new technologies supplant processes formerly handled by people, two contradictory but linked sentiments arise. We celebrate the efficiency the technology affords us, but also fear that knowledge and skills will atrophy. Spellcheck might mean our error rates decrease and we do not have to dedicate the same cognitive load to a menial evaluation task, but if we are not reallocating our load to handle more global planning work, and we cannot easily reclaim the spellcheck process, then is the total cognitive load, our intelligence, shrinking?

This threat is difficult to evaluate with an act as complex as writing. Because no one technology is likely to "solve" for writing, an entire corpus of technologies continues to chisel away at and add to it. But how do we know we are preserving the essential elements? In his Phaedrus dialogue, Plato critiques writing, then a new technology, on the grounds that it "will create forgetfulness in the learners' souls, because they will not use their memories," an older technology [44]. Considering the plasticity of the brain, Plato was right to fear that externalizing memory and experience into writing would erode certain faculties, but he had no way of knowing the world made possible by writing, a world that writing-natives would not give up even for a preternatural memory. How we define and value a technology is not inviolate and immutable, but instead changes to reflect the biases of whoever happens to employ that technology.

To understand which processes should stay in the hands of the writer (for now) and which can be decomposed and sourced, we should not consider how these new technologies can alter writing, but how deeply our understanding of the world is founded in to the act of writing. Writing has a rich history and is entwined with our understanding of language and communication, so even if new technologies can change the writing process overnight, our cultural values around writing will be slower to change. Though we feel okay sourcing a "menial" task like spellchecking to a machine, we are probably not ready to offload a task like word choice, which carries more meaning.

In order to preserve writing skills we are not ready to jettison, and smooth the transition from a more solitary, personal writing process to a more collaborative and computer-mediated writing process, we believe that writing tools can improve writing not just by making the process more efficient or by improving the end product, but by teaching people to be better writers. Research has shown a number of principles are essential to the learning process. For one, writing students do not learn as effectively when their errors are cleaned up automatically. Asking writers to "*notice* the gap" between their own writing and corrected versions of that writing helps them bridge between the two states. Noticing is important in transforming explicit knowledge (i.e., an understanding of rules) into implicit ability (i.e., strategies that successfully employ those rules [51]). Additionally, writing always has a real-life purpose, so the best learning occurs in situ, not as exercises separated from *context*. Sometimes skills learned in drills do not translate to other contexts and therefore a diversity of contexts is also valuable. Similarly, rule-based drilling is less effective than allowing a writer to actively *transform* text. Explicit, rule-based instruction is helpful, but only when sequenced after a writer has been made to actively engage with the text [15]. Combining shorter, simpler syntactic units into larger, more complex syntactic units is especially helpful [19].

To illustrate these concepts is an example decomposition, recomposition, and sourcing workflow that sources much of the work away from the self, but promotes learning for all three labor sources:

1. *Machine* identifies sentences that seem awkward,
2. *Self* selects sentence to send to *crowd* (noticing),
3. *Machine* recruits *crowd* to participate in a sequence of tasks: Highlight, Mutate, See (context),
4. *Machine* combines sentence with surrounding sentences, sends to *crowd* many times over (context),
5. *Crowd* (transforms) sentence to improve readability,
6. *Machine* chooses the three most common fixes, sends back to *self* and to *crowd*,
7. *Crowd* sees their own transformation against group selections with percentage frequency (noticing),
8. *Self* selects transformations blind and after selection has group selection percentages revealed (noticing),
9. *Machine* notes *self*'s selection to inform future Step 1s (noticing, context), and
10. *Machine* notes *crowd*'s selection against consensus & *self*'s selection to inform future Step 4s (noticing, context).

Step 8 is not time or energy efficient and may seem counterintuitive, but carries potential benefits. While the self

has sourced all microtasks that require transforming text, the self still gets to notice the gap in a contextually-rich environment. Additionally, future machine steps will be informed by the self's selection, potentially reducing time and effort in the long run. It is this long-run view that we advocate researchers and technology developers adopt.

CONCLUSION

We have provided a brief introduction motivating the development of advanced technological support for writing, beyond traditional tools that mostly support formatting and organization. Building tools to support writing will be challenging because writing is a complex, networked process that involves a variety of task types, but it will also be rewarding as writing is a fundamental but poorly-understood skill. We believe that developing tools will provide a valuable lens by which to understand and explore interesting fundamental problems related to a number of domains, including human-computer interaction, composition studies, and natural language processing.

ACKNOWLEDGMENTS

This report benefited from the support of the Microsoft Research Selfsourcing Team, including Ece Kamar, Dan Liebling, Kristina Toutanova, Saleema Amershi, and Andres Monroy-Hernandez, and interns Carrie Cai (MIT), Rhema Linder (Texas A&M), Niloufar Salehi (Stanford), Ke Tran (University of Amsterdam), and Rajan Vaish (UCSC).

REFERENCES

1. Ackerman, D. S., & Gross, B. L. (2005). My instructor made me do it: Task characteristics of procrastination. *Journal of Marketing Education,* 27(1): 5-13.
2. Agapie, E., Monroy-Hernandez, A., & Teevan, J. (2015). Crowdsourcing in the field: A case study using local crowds for event reporting. In *Proceedings of HCOMP*.
3. Argenti, P. A. (2006). How technology has influenced the field of corporate communication. *Journal of Business and Technical Communication*, *20*(3).
4. Arum, R., & Roksa, J. (2011). *Academically Adrift: Limited Learning on College Campuses.* University of Chicago Press.
5. Bangert-Drowns, R. L. (1993). The word processor as an instructional tool: A meta-analysis of word processing in writing instruction. *Review of Educational Research*, 63(1).
6. Bernstein, M. S., Brandt, J., Miller, R. C., & Karger, D. R. (2011). Crowds in two seconds: Enabling realtime crowd-powered interfaces. In *Proceedings of UIST*.
7. Bernstein, M. S., Little, G., Miller, R. C., Hartmann, B., Ackerman, M. S., Karger, D. R., Crowell, D., & Panovich, K. (2010). Soylent: A word processor with a crowd inside. In *Proceedings of UIST*.
8. Birnholtz, J., & Ibara, S. (2012). Tracking changes in collaborative writing: Edits, visibility and group maintenance. In *Proceedings of CSCW*.
9. Cai, C., Iqbal, S.T., & Teevan, J. (2016). Chain reactions: The impact of order on microtask chains. In *Proceedings of CHI*.
10. Casner-Lotto, J., & Barrington, L. (2006). *Are They Really Ready to Work? Employers' Perspectives on the Basic Knowledge and Applied Skills of New Entrants to the 21st Century US Workforce*. Partnership for 21st Century Skills.
11. Cheng, J., Teevan, J., Iqbal, S.T., & Bernstein, M.S. (2015). Break it down: A comparison of macro-and microtasks. In *Proceedings of CHI*.
12. Chilton, L., Little, G., Edge, D., Weld, D.S., & Landay, J. (2013). Cascade: Crowdsourcing taxonomy creation. In *Proceedings of CHI.*
13. Dow, S., Kulkarni, A., Klemmer, S., & Hartmann, B. (2012). Shepherding the crowd yields better work. In *Proceedings of the CSCW*.
14. Dvorak, J.C. (2004). Top 10 most important software programs. *PC Magazine*, August 2004.
15. Farr, M., & Daniels, H. (1986). Language Diversity and Writing Instruction. National Council of Teachers of English.
16. Ferrari, J.R., Johnson, J.L., & McCown, W.G. (1995). *Procrastination and Task Avoidance: Theory, Research, and Treatment*. Springer Science & Business Media.
17. Flower, L. (1979). Writer-based prose: A cognitive basis for problems in writing. *College English*, 19-37.
18. Flower, L., & Hayes, J.R. (1981). A cognitive process theory of writing. *College Composition and Communication*, 365-387.
19. Graham, S., & Perin, D. (2007). A meta-analysis of writing instruction for adolescent students. *Journal of Educational Psychology*, 99(3).
20. Gray, F.E., Emerson, L., & MacKay, B. (2005). Meeting the demands of the workplace: science students and written skills. *Journal of Science Education and Technology*, 14(4).
21. Hayes, J.R., & Flower, L. (1981). Uncovering Cognitive Processes in Writing: An Introduction to Protocol Analysis.
22. Ipeirotis, P. (2013). Answer to "How many active workers are there in Mechanical Turk?" Quora, retrieved December 2015.
23. Kellogg, R.T. (1988). Attentional overload and writing performance: Effects of rough draft and outline

strategies. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14(2).

24. Kidd, C., Piantadosi, S.T., & Aslin, R.N. (2012). The Goldilocks effect: Human infants allocate attention to visual sequences that are neither too simple nor too complex. *PloS one*, *7*(5), e36399.
25. Kim, J., Cheng, J., & Bernstein, M.S. (2014). Ensemble: Exploring complementary strengths of leaders and crowds in creative collaboration. In *Proceedings of CHI*.
26. Lasecki, W.S., Kamar, E., & Bohus, D. (2013). Conversations in the crowd: Collecting data for task-oriented dialog learning. In *Proceedings of AAAI*.
27. Lasecki, W.S., Miller, C., Sadilek, A., Abumoussa, A., Borrello, D., Kushalnagar, R., & Bigham, J. (2012). Real-time captioning by groups of non-experts. In *Proceedings of UIST*.
28. Lasecki, W.S., Wesley, R., Nichols, J., Kulkarni, A., Allen, J.F., & Bigham, J.P. (2013). Chorus: A crowd-powered conversational assistant. In *Proc. of UIST*.
29. Marzano, R.J. (2007). *The Art and Science of Teaching: A Comprehensive Framework for Effective Teaching*. Alexandria, VA: ASCD.
30. Marzano, R.J. (2009). *The Art and Science of Teaching/Helping Students Process Information*.
31. Montuori, A., & Purser, R.E. (1995). Deconstructing the lone genius myth: Toward a contextual view of creativity. *Journal of Humanistic Psychology*, 35(3).
32. Mullen, T. & Collier, N. (2004). Sentiment analysis using support vector machines with diverse information sources. In *Proceedings of EMNLP*.
33. National Center for Education Statistics (2012). *The Nation's Report Card: Writing 2011* (NCES 2012–470). Institute of Education Sciences, U.S. Department of Education, Washington, D.C.
34. National Commission on Writing. (2004). Writing: A ticket to work… Or a ticket out? A survey of business leaders.
35. Nebeling, M., Guo, A., Murray, K., Tostengard, A., Mihaijlov, M., Dow, S., Teevan, J., & Bigham, J. (2015). WearWrite: Orchestrating the crowd to complete complex tasks from wearables (we wrote this paper on a watch). arXiv:1508.02982.
36. Nebeling, M., To, A., Guo, A., de Freitas, A., Teevan, J., Dow, S., & Bigham, J. (2016). WearWrite: Crowd-Assisted writing from smartwatches. In *Proc. of CHI*.
37. Neuwirth, C.M., Kaufer, D.S., Chandhok R., & Morris, J. H. (2000). Computer support for distributed collaborative writing: A coordination science perspective. *Coordination Theory and Collaboration Technology*.
38. Noël, S., & Robert, J.M. (2004). Empirical study on collaborative writing: What do co-authors do, use, and like? *Computer Supported Cooperative Work* 13(1).
39. Nunes, M.J. (2013). The Five-Paragraph Essay: Its Evolution and Roots in Theme-Writing. *Rhetoric Review*, 32(3).
40. Oatley, K., & Djikic, M. (2008). Writing as thinking. *Review of General Psychology*, 12(1).
41. Olive, T. (2004). Working memory in writing: Empirical evidence from the dual-task technique. *European Psychologist*, *9*(1).
42. Perl, S. (1979). The composing processes of unskilled college writers. *Research in the Teaching of English*.
43. Pinker, S. (2014). The source of bad writing. *The Wall Street Journal*, September 25.
44. Plato, & Stephen Scully. 2003. *Plato's Phaedrus*. Newburyport, MA: Focus Pub./R. Pullins Co.
45. Posner, I., & Baecker, R. (1992). How people write together. In *Proceedings of HICS*.
46. Shah, C., Erhard, K., Ortheil, H.J., Kaza, E., Kessler, C., & Lotze, M. (2013). Neural correlates of creative writing: an fMRI study. *Human Brain Mapping*, *34*(5).
47. Scardamalia, M., & Bereiter, C. (1987). Knowledge telling and knowledge transforming in written composition. *Advances in Applied Psycholinguistics*, 2.
48. Tam, J., & Greenberg, S. (2006). A framework for asynchronous change awareness in collaborative documents and workspaces. *International Journal of Human-Computer Studies*.
49. Teevan, J., Iqbal, S.T., & von Veh, C. (2016). Supporting collaborative writing with microtasks. In *Proceedings of CHI*.
50. Teevan, J., Liebling, D., & Lasecki, W.S. (2014). Selfsourcing personal tasks. In *Proceedings of CHI*.
51. Thornbury, S. (1997). Reformulation and reconstruction: Tasks that promote 'noticing'. *ELT Journal*, 51(4).
52. Toutanova, K., Klein, D., Manning, C.D., & Singer, Y. (2003). Feature-rich part-of-speech tagging with a cyclic dependency network. In *Proc. of NAACL HLT*.
53. Wang, J., Faridani, S., & Ipeirotis, P. (2011). Estimating the completion time of crowdsourced tasks using survival analysis models. In *Proc. of CSDM*.
54. Yen, T.H., Wu, J.C., Boisson, J., Chang, J., & Chang, J. (2015). WriteAhead: Mining grammar patterns in corpora for assisted writing. *ACL-IJCNLP*.

APPENDIX A: THE RHETORICAL SITUATION

The rhetorical situation is a useful way of thinking and talking about writing and texts, and important to understand when building technological support for writing. The rhetorical situation encompasses aspects of the context of the writing, including the message, writer and audience, as represented in Figure A.1. Each aspect is discussed in greater detail below.

Figure A.1. The rhetorical situation is often drawn as a triangle circumscribed in a circle, to represent three interacting agents (the message, the writer, and the audience) in a larger medium.

Message

The message is the purpose of the text. It addresses what the writer trying to say and why the writer trying to say it. Sometimes there are layers of messages. A presidential state of the union address, for example, exists to communicate specific details about government function to the larger public, but it also exists to make the public feel safe. These two messages are layered and might interact with one another. Saying, "Our country has $18.628 trillion in debt," might be an accurate statement, but might scare the public more than a phrase like "debt is only up by the less than .1% from last year."

Message is not just content, but form too. The way in which we communicate is in itself a message. This is especially evident with newer media and technologies like e-mail, which have their own conventions and idiosyncrasies (e.g., timestamps, signatures, BCC fields, etc.). A good writing technology does not just help writers produce prose in a bubble, but in response to its form.

Writer

The writer (or speaker) is the author of the message, but the writer is more complex than a single person with a writing tool. A writer might be many individual people, as in a presidential address, which is written by a team of speechwriters but delivered by a single voice. A writer implicitly represents other groups—demographics, institutions, vocations—but a writer can also inhabit a persona to highlight certain attributes of the self. As such, the way the message is delivered constructs the audience's understanding of the writer and in turn the meaning of the message. If the president is too sly with word choice and appears to be masking the aforementioned debt, this could lead an audience to misinterpret or outright reject the intended message.

Writers can be unaware of how they come across, which is a problem in business communication, which often requires navigating power differences. Writers also tend to rely on repetitive syntax, idioms, or writing strategies, which might not be appropriated for their message. Technology can identify these kinds of habits for writers for might not be aware of them.

Audience

The audience is the recipient of the message, and like the writer, the audience can be one or many people, literal or abstract. The expectations of the audience determine the content and form of a successful message (i.e., one that fulfills its purpose). Writers must consider the power relations and history between themselves and their audience. Broader audiences will not necessarily have the technical background to understand messages from the scientific or academic community and so messages will have to be distilled to the basics or translated into layman's terms. A good writing technology will be able to identify if and when writing fits the audience's expectations or knowledge level.

Context

The context is the set of larger conditions surrounding the writer, audience, and message. These factors are typically atmospheric and not readily constructed by the message or the writer, which is why they are represented as the space in which the situation occurs rather than the situation itself. Context may be subdivided into different subcategories. The location is the physical or abstract space in which a rhetorical situation occurs. A presidential address has an immediate physical setting with an immediate physical audience, but when encoded in the medium of a video broadcast, its location expands into a virtual space. A similar translation occurs with time in that a message can be delivered asynchronously by programming it in a broadcasting medium. Messages also occur in a specific culture or confluence of cultures. A presidential address considered patriotic and rousing in a home country might appear militant or jingoistic in another. Lastly, a genre is when artistic features of a form become codified through repetition. In a presidential address, affixing an American flag pin to the left lapel is such an automatic expectation that presidents who have opted not to wear it have been criticized for being unpatriotic.

Context is especially important in business writing. A significant amount of business communication occurs in IM and via text messaging, but the informal conventions of these forms (e.g., improper punctuation) are not necessarily appropriate for an email to a superior or a written letter to a client.