Conceptual Approach for the Design of Radiology Reporting Interfaces: The Talking Template

Chris L. Siström, M.D.

Within the coming decade, traditional dictation supported by human transcription for radiology reports will be replaced by one or more computerized methods. This paper discusses the cognitive and process efficiency problems arising from currently available technology including speech recognition and menu-driven interfaces. A specific concept for interaction with the reporting interface is proposed. This is called the "talking template" and departs from other designs by providing for all interactions to be mediated through audible prompts and microphone controls. The radiologist can recapture efficiency and cognitive focus by dictating while viewing images without the "look away" problem inherent in other interfaces.

KEY WORDS: Radiology reporting, user interface, process efficiency, cognitive function, talking template

INTRODUCTION

A recent paper outlined a general framework for improved radiology reporting. In it, three report attributes were articulated. These were structured format, consistent content, and standard language.¹ Herein, I will discuss two additional aspects of radiology reporting that are manifest during document preparation by radiologists. These are reporting efficiency and cognitive process. I will make the case that traditional dictation supported by human transcription is still the most efficient method for the radiologist. Furthermore, dictating while simultaneously searching images may have intrinsic merit by virtue of cognitive synergy between the two tasks. I will assert that both of the emerging technologies—speech recognition and menu-oriented methods—as currently implemented, force radiologists to disaggregate the two tasks resulting in inefficiency and cognitive overload. The proposed solution avoids a

visual interface in favor of interaction with the reporting system by audible prompts and microphone controls. A suggested term for this class of reporting interfaces is the "talking template." The purpose of this paper is to present an argument for using the talking template concept (or something very much like it) as the reporting interface for digital dictation systems seeking to replace traditional dictation supported by human transcription. This design approach makes no assumptions about specific technologic details, vendors, or distribution models (open source or commercial).

THE COGNITIVE PROCESS OF REPORTING

One way to classify the cognitive task of radiology reporting relates to the relative number of separate images and sequences contained in the examination under review. At the low end of this spectrum would be two views of the chest with no comparison. Here, the entire study is displayed simultaneously and little, if any, manipulation of the images is required during their evaluation. The radiologist can look at the images, formulate what needs to be in the report, and then turn away and speak, type, or do menu selections to record the

From the Department of Radiology, University of Florida School of Medicine, PO Box 100374, Gainesville, FL 32610, USA.

Correspondence to: Chris L. Siström, M.D., Department of Radiology, University of Florida School of Medicine, PO Box 100374, Gainesville, FL 32610, USA; tel: 352-395-0291; fax: 352-395-0279; e-mail: sistrcl@radiology.ufl.edu

Copyright © 2005 by SCAR (Society for Computer Applications in Radiology)

*Online publication 2 June 2005
doi: 10.1007/s10278-005-5167-8*

findings. If memory must be refreshed, a single glance back to the images will suffice. At the other end of the spectrum would be a multisequence MRI or CT examination. In these latter examples, continuous use of the imaging workstation display tools is required to view the entirety of the imaged volume. For examinations such as CT angiography, proper interpretation requires performing rather involved post processing of the original images. Therefore, while not impossible, the strategy of looking at the images first, formulating the interpretation mentally, and then creating the report document is much more difficult. If the radiologist wishes to look back to the images, he or she would likely need to manipulate them to return to the anatomic location and presentation state desired.

A second conceptual division in radiology interpretation has to do with the form of clinical question being answered by means of the report. In some types of study, there is an unambiguous question about a particular disease. This might be called a closed form question. The answer is often dichotomous (disease positive, disease negative) although it might be ordinal or quantitative (probability of disease). The classic example is disease surveillance as in screening mammography, where the major question is "rule out breast cancer." Slightly more complex examinations such as routine carotid ultrasound still have a single major underlying clinical concern (e.g., look for significant atherosclerotic disease). At the other end of the clinical question continuum are survey examinations that are symptom- or problem-driven. These might be termed open-ended clinical questions. Here, there are multiple possible findings that each may represent one or more various disorders that must be considered and addressed. Details of the clinical situation will interact with complex imaging findings while formulating a diagnostic conclusion. An example of this sort of more complex interpretive task would be a CT scan of chest, abdomen, and pelvis performed three days after liver transplant on a patient with fevers and failure to wean from mechanical ventilation.

Much of the early work in computer-assisted reporting was carried out to facilitate radiology examinations that would be considered cognitively simple (although by no means easy) by the framework outlined above. Again, screening mammography serves as perhaps the best example with relatively few images, a single clinical question

having definable answers, and limited types of findings (calcifications and soft tissue abnormalities). Other examinations that have served as substrates for computer-mediated report generation include orthopedic trauma, obstetrical ultrasound, and echocardiography. Workers at Johns Hopkins,²⁻⁵ the University of California,⁶⁻⁸ Harvard,⁹⁻¹¹ and others^{12,13} have developed and described such systems. These also tend toward the simpler end of the cognitive spectrum. As clinical imaging data sets become increasingly large and complex, much of the radiologist's time and cognitive effort must be devoted to manipulating the display and postprocessing controls of the imaging workstation. The notion of introducing a second, equally complex interface for report generation through semantic navigation raises a critical question. How will the radiologist work with a complex reporting interface without losing efficiency and/or cognitive focus on the images?

The great advantage of traditional dictation methods is the complete lack of what has been called the "look away" problem inherent in many of the alternative technologies. This includes both speech recognition systems (SRS) as they are typically designed and used, as well as so-called point and click (menu-driven) interfaces. During traditional dictation, all interactions with the system are performed via the dictation microphone. The microphone, integrated navigation buttons, and a built-in playback speaker facilitate these interactions. Thus, the radiologist's eyes never leave the images while dictating and his other hand is free to manipulate the image display controls. In addition to being maximally efficient, the traditional arrangement may actually enhance cognitive focus because the radiologist simultaneously searches the images and describes their findings. Although there is ample research into eye gaze patterns during interpretation of images by radiologists, there is little, if any, empiric research that provides an estimate of how speaking synergizes with looking at radiology images.¹⁴⁻¹⁶ Many radiologists can recall instances of dictating something like "The bones are..." and noticing the thoracic compression fracture at that moment. Readers are encouraged to think of relevant examples of this phenomenon from their own experience. Is there some inherent cognitive advantage to dictating while viewing the images that might be lost by switching to alternative methods?

This issue has not been addressed formally in the literature about improving radiology reporting.

There is a sizable body of research that shows that humans are able to use and integrate input from different sensory modalities when performing various cognitive tasks.^{17–20} In dual-mode cognitive processing, some combination of sensory inputs reinforce each other. For example, seeing a series of printed words and simultaneously hearing them spoken allows most people to remember a larger number and for a longer time. Another example can be found in the mnemonic used by some politicians to remember names of the many people that they meet. It involves deliberately making some visual association with the person and their name. Similar examples abound in the neuropsychological literature. One reason for the intense interest in dual-mode sensory processing is the theory that dyslexia is related to modal disconnection. The hypothesis is that dyslexics have difficulty fusing input from separate senses.^{21–23} Thus, their cognitive performance may actually suffer when presented with multiple complex inputs. Are we risking a form of induced dyslexia in radiologists by artificially disconnecting the formerly complementary cognitive tasks of viewing images and dictating the findings?

Aside from the disadvantage of having to look away from the images during reporting, there are two additional cognitive problems with reporting methods using computer menus rather than dictated narrative text. These have to do with how clinicians understand and communicate complex medical information. Narrative written or verbal communication is certainly what clinicians are trained in and accustomed to using. There is emerging consensus that attempting to “computerize” such cognitive processes may be detrimental. In a recent paper about healthcare information technology, Ash cites evidence from studies in cognitive psychology and sociology that in a shared context, concise, unconstrained, and free-text communication is the most effective way for coordinating work around a complex task.^{24,25} Overly structured data can lead to loss of cognitive focus by clinicians, both during input and review.²⁶ This can cause clinicians to experience a loss of overview about the case at hand when they have to attend to data contained in many different fields, sometimes on different screens within a graphical interface.^{27,28} Furthermore, the act of writing or

dictating in narrative form may be integral to the cognitive processing of the case.²⁹

REPORTING EFFICIENCY

Most radiologists who have experience reporting with traditional dictation supported by a transcriptionist and clerical personnel to hang films on a multiviewer will agree that this is the most efficient method from their perspective. Note that the efficiency of the entire process may not be ideal in this situation because the time between dictation and report availability is variable and relatively long on average. Moreover, because radiologist review of the report documents prior to signature happens hours to days after the dictation, specific details of each case might not be fully remembered. This could result in errors of transcription going undetected or being incorrectly addressed. To remedy this problem requires “rework” on the part of the radiologist to review the images, which can be cumbersome and time-consuming even in a soft copy reading environment. After introduction of SRS into clinical radiology practices, radiologists virtually are unanimous in perceiving that their personal efficiency is decreased. Estimates of the increase in time required per case range from 50% to 100%.^{30–34} Most radiologists recognize that the large improvement in report turnaround time is worth some additional effort on their part.³⁵ However, increasing volumes of imaging workload per physician exerts countervailing pressure to increase individual productivity. Extensive use of complete and partial predefined reports as well as templates may serve to reclaim some of the lost efficiency in SRS although there is little empiric evidence about this.³⁶

There are two reasons why the traditional method of dictation followed by human transcription is efficient. Both are results of the fact that the radiologist—by definition—cannot check the transcription during the reporting session because it occurs sometime afterward. The first efficiency is because the dictation part of the task is necessarily done in a single step for each case. The second efficiency is directly related to the first. Because the radiologist cannot check the transcription of their dictation, he or she has no choice but to move on to the next study after finishing the current study.

Thus, examinations are dictated in batches and then later edited and corrected all at once. The concept of having workers repeatedly perform a single part of a multistep process is central to modern industrial theory. By analogy, batch mode dictation is more efficient than serially performing every step in the reporting process for each case. With this in mind, it is entirely appropriate that modern SRS systems were designed specifically to support established work patterns of radiologists (i.e., batch mode dictation followed by editing a group of reports). However, most radiologists learn to use SRS in a counterintuitive way whereby they interact intensely with the graphical SRS interface to produce reports singly rather than in groups or batches. This introduces modest inefficiency when performing low-volume, high-complexity studies because the report production is a relatively small part of the overall cognitive effort required for the task. However, for high volume work (e.g., intensive care chest films), the extra time spent creating report documents one at a time can far exceed the time employed in looking at the images. This has been termed the “look away” problem and it is shared by SRS and other computer mediated reporting methods. This trend is exemplified in over 6 years of experience with SRS in our institution.

The Radiology Department at The University of Florida switched from transcriptionist-supported reporting to SRS for all studies over a single weekend in 1998. Subsequently, I have observed over 50 residents and fellows learn to use SRS as their first and only method for producing reports. An additional 35 faculty members switched to SRS after using traditional dictation/transcription methods. During that entire time, only one radiologist in the practice (myself) has regularly used “batch mode” dictation for high-volume repetitive work. Otherwise, radiologists universally dictate one examination at a time, edit, and verify or sign it before moving on to the next case. I have asked many trainees and staff radiologists about this and the answer most often has to do with concern about not remembering the details of the case if editing the report is deferred. The irony of this belief is that with human transcription, review and editing of the report would be performed much longer after dictation than with SRS. Another observation concerns how radiologists interact with the SRS interface. There seems to be fascination bordering

on obsession with checking to see if the recognition of the last couple of sentences was accurate. When errors are spotted, radiologists seem to be compelled to correct them immediately. Thus, the user is attending to the SRS interface during much of the time taken to create the report rather than viewing the relevant images while dictating. In many instances, radiologists frequently shift attention back and forth between the images and the SRS interface during reporting. Thus, both types of efficiency described above are completely lost.

Many radiologists, faced with the frustrations of SRS, seem to find that using predefined templates or macros helps them to regain efficiency or at least some level of confidence in and control over the process.^{35–37} With voice-activated and predefined phrases, headings, and paragraphs, one can be sure that there will be no errors in their content although activation of the correct predefined text is imperfect. This means that the radiologist is again compelled to attend closely to the editing interface to move through the report and to check that the predefined content is properly inserted. The same sort of inefficiency is introduced with reporting systems that require the radiologist to navigate through some sort of visual interface to select findings for inclusion in the report. These are the so-called point-and-click interfaces. By definition, if the user is attending to an interface to locate and activate appropriate findings, he or she is not looking at the images (i.e., look away time). With respect to the process of reporting, there are other potential problems with menu-driven, forced choice systems. These will be outlined below along with the hoped for advantages of this approach.

TEMPLATE-DRIVEN REPORTING

Although not formalized or explicit, templates are employed routinely by radiologists using traditional dictation with human transcription. These templates are internalized, learned during training, and reinforced through experience. Almost all radiologists use at least a rudimentary structure in their reports. In the simplest form this includes: indication, prior studies, procedure details, findings, and impression. Additionally, they generally have mental “check lists” (templates) relevant to various examination and indication combinations that serve to guide the search for abnormalities.

Traditional dictation allows the interpreting physician to simultaneously bring up the next item on their internal template, search the image(s) for relevant findings, and dictate their observations. This method is not only maximally efficient but also takes advantage of the power of multimodal cognitive processing as described above. One could argue that a large part of traditional training of radiology residents is to provide each of them with a personal set of mental checklists or templates for use during their own practices.³⁸

Given a natural predilection on the part of radiologists to use internal “check lists” during reporting, it makes intuitive sense to consider shared and formal templates among groups of radiologists. Indeed, many practices have done this sort of thing for decades by means of paper-based forms or “cheat sheets” posted at reporting stations. Commercially available SRS systems have variably robust facilities for creating and navigating through templates or report shells. One common method is to define stopping points (delimited with brackets or other special characters) within the template. Often, these are the traditional sections of a report—indication, prior studies, and so on—with added headings within the findings section. Example of such templates are included in Exhibits 1 and 2. Using voice

Exhibit 1. Sample report template for unenhanced head CT scan.

UNENHANCED COMPUTED TOMOGRAPHY OF THE HEAD
COMPARISON: [dictation point]
PROCEDURE: [dictation point]
IMAGE QUALITY: [dictation point]
BONES-SINUSES: [dictation point]
EXTRACRANIAL SOFT TISSUES: [dictation point]
VENTRICLES: [dictation point]
BRAIN-DEVELOPMENTAL: [dictation point]
BRAIN-AGE RELATED: [dictation point]
BRAIN-POSTIVE FINDINGS: [dictation point]
BRAIN-NEGATIVES: [dictation point]
EXTRA-AXIAL SPACES: [dictation point]
VASCULAR STRUCTURES: [dictation point]
OTHER FINDINGS-COMMENTS: [dictation point]
IMPRESSION: [dictation point]

Exhibit 2. Sample report template for abdominal (±pelvis) CT scan.

ABDOMINAL COMPUTED TOMOGRAPHY
COMPARISON: [dictation point]
PROCEDURE: [dictation point]
LIVER: [dictation point]
GALLBLADDER/BILIARY: [dictation point]
PANCREAS: [dictation point]
SPLEEN: [dictation point]
KIDNEYS/URETERS: [dictation point]
ADRENALS: [dictation point]
AORTA/VESSELS: [dictation point]
BOWEL/APPENDIX: [dictation point]
FLUID/FREE AIR: [dictation point]
LYMPH NODES: [dictation point]
PELVIC ORGANS: [dictation point]
OTHER FINDINGS: [dictation point]
IMPRESSION: [dictation point]

commands or a microphone button, the radiologist can advance through the stopping points and dictate relevant narrative text in each section. This method is popular with many radiologists faced with having to use SRS for the first time. However, this necessitates considerable “look away” from the images to make sure that one is dictating under the correct heading.

In addition to being in concert with radiologist’s natural way of working, templates have other advantages. If dictation shells are shared and standardized, the resulting reports will have similar content and common ordering. By extension, interpretations will be more complete and consistent because standard templates serve as checklists for the authors. There is a large body of published research detailing the opinion of referring physicians regarding the content and format of radiology reports.^{39–44} When surveyed, referring clinicians express strong preference for the reports that they receive to be in the kind of itemized, structured, and complete format encouraged by the use of templates. Interestingly, in experiments comparing free text to structured format, senior medical students are equally accurate and efficient at reading radiology reports for specific content.⁴⁵ As with referring clinicians, these subjects strongly preferred the structured format to free text.

There is one more advantage to using predefined shells for reports that include headings or labels in

the final document. Consider how a radiologist would have to construct a narrative sentence to express that the liver is normal in dictating a cross-sectional imaging examination. He or she might say; "New paragraph, The liver is normal" or "New paragraph, There are no liver masses or other abnormalities." In a report shell or template that already contains a heading called LIVER, it is sufficient to simply dictate "negative" after the heading. I use the term telegraphic speech to describe this type of dictation. In practice, if one can master navigation to the appropriate headings, telegraphic speech is more efficient than constructing entire sentences or paragraphs to express the same things. Motivated by this and the preceding reasons, I assert that any proposed technical solution for achieving improved reporting with structured format and consistent content will require some form of report template to be integral to the design. However, as currently implemented, these report shells are presented as text on an editing screen. Thus, visually navigating through the template gets us back to the thorny issue of "look away" time described above.

MENU-DRIVEN REPORTING

In light of the discussion above about the value of templates, it is ironic that many computerized, menu-driven reporting systems are essentially passive. This means that the radiologist must first decide exactly what he or she wishes to report, and in what order, before attempting to find it in the menu structure of the reporting interface. To give appropriate credit to the designers of these systems, there is often rather sophisticated logic to aid the radiologist once a class of finding is chosen. For example, selecting a high level menu item often triggers relevant choices for describing the particulars of the abnormality or structure in question. The passive nature of menu-driven systems could conceivably be overcome by adding some form of meta-logic to the system that would encode templates for particular report types. Kahn has developed an innovative technical framework for doing this.⁴⁶⁻⁵² The system, called SPIDER, employs self-defining report documents using extensible markup language (XML). Within these dynamic documents, both the template structure headings and the relevant categorical choices for

each one are completely defined. These documents are designed to be completed with a simple web browser, keyboard, and mouse. It is easy to imagine how they could be "filled out" using voice commands and speech recognition for text blocks. Technology very similar to this is currently available in the latest version of at least one major vendor's SRS offering. The concept of using a sophisticated menu-driven interface with predefined report shells that provide consistent structure to the report is quite attractive. However, the very sophistication of the concept means that the interface will almost certainly need to be rather complex both cognitively and visually. At the very least, the radiologist would see a series of headings and would have to navigate to each one in turn so as to select the appropriate choice and/or dictate supplemental text. Such a complex and demanding interface would require considerable attention by the radiologist and introduces yet again the seemingly intractable "look away" problem.

One necessary artifact of a menu-driven interface is that it causes the radiologist to respond to multiple "forced choice" situations. Designers of the lexicon and the ontology are faced with a Hobson's choice concerning how many categories to provide for a given conceptual unit. Parsimony requires that the number of categorical levels be limited as much as possible. However, accuracy and precision often call for a large number of levels. This problem is well known to workers in diverse areas of medicine. Researchers designing survey instruments for a multitude of purposes, developers of clinical coding systems for computerized patient records, and those responsible for defining disease staging systems all share the same dilemma. With an insufficient number of levels, some individual observations will be ambiguously assigned. With too many choices, the system becomes cumbersome to communicate and use in routine practice. Consider a radiologist faced with codifying a lung mass. First, he or she must characterize the mass in their own mind and then turn to the menu and scan the available choices. They will initially try and find a choice set that fits the current mass because shifting to typing or free-text dictation represents an extra step. If too few choices are available in the menu, ambiguous or intermediate observations may be misclassified by fitting them to the nearest matching choice. If too many categories are available,

scanning them is inherently time-consuming. This example illustrates the problem mentioned above, that forcing clinicians to code their work in overly structured ways can lead to cognitive overload, frustration, and errors.²⁶⁻²⁸

If menu-driven, forced-choice radiology reporting interfaces are problematic in terms of cognitive overload and inefficiency of process, why would we want to use them? The rationale often advanced is that report documents in codified form will lend themselves to several admirable enterprises. Among these are to enhance quality assurance and process improvement efforts within departments and hospitals. Another goal is to codify report content to meet the needs of billing and compliance efforts within the organization. Perhaps the most compelling purpose for codified report content is to enhance research efforts by eliminating the need for labor-intensive abstraction of narrative interpretative documents. Note that none of these goals relates to the basic purpose of the radiology report: clinical communication between providers during the actual provision of healthcare. With this in mind, I would suggest that any technological methods for improving reporting must respect and acknowledge this primary reason for engaging in the process. The needs of researchers, administrators, informatics workers, and bureaucrats must be accommodated only to the extent that they do not interfere with patient care. When thinking about the design of radiology reporting systems, facilitating and recording efficient, high fidelity, and transparent communication between the radiologist and referring physician is the primary goal. All of the other considerations are secondary and should be put aside if they interfere with the basic function of clinical communication. As I will assert below, it should be possible to optimize the process of radiology reporting by combining efficient use of templates, speech recognition, and speech synthesis into the proposed "talking template" interface. The secondary goals can then be met by use of natural language processing (NLP) on the resulting semistructured report documents.

There is an important and clinically relevant reason to consider a forced choice response for at least one element of the report. That is in the conclusion or impression section. Radiology trainees are taught to be definitive and to make diagnostic commitments in their reports. However,

when faced with complex, ambiguous, or indeterminate findings, practicing radiologists are tempted to simply list the findings and not commit to a diagnostic conclusion. A perfect example of this problem and its resolution can be found in mammography. A mammogram report that simply describes masses and/or calcifications and does not come to a definitive recommendation is practically useless to the clinician and their patient. The Breast imaging reporting and data system (BIRADS), promulgated by the ACR and now partly codified in the MQSA law, requires a definitive diagnostic conclusion.⁵³ This is the familiar six-point scale consisting of; need additional imaging (A), negative (N), benign (B), probably benign (P), suspicious (S), and malignant (M). This forced choice actually simplifies the cognitive task of reporting a mammogram and has been shown to increase reproducibility.^{54,55} Readers of BIRADS coded reports now know exactly what they need to do with the information because the follow-up recommendations are integral to each result category. A similar forced choice system is used in Veteran's Administration Hospitals for coding all radiology reports with an ordinal scale relating to the relative abnormality and acuity of the examination findings. Such scales may enhance clinical communication and reduce errors by encouraging direct communication with referring clinicians for unexpected or high acuity findings.

NATURAL LANGUAGE PROCESSING

Although natural language processing (NLP) may sound like futurist speculation, related technology is already employed in currently and widely used SRS systems. All versions of commercially available SRS software use rather sophisticated semantic analysis and statistical modeling to choose the most likely words and phrases matching the speech input. One of the trickiest design issues with modern SRS systems is that they are essentially context-free. By this, I mean that the system must be able to accept and translate free-text narrative dictation within the entire domain of radiology interpretation. No assumptions are made about the type of examination being interpreted, the body area involved, or the organ system being described. Consider how much easier the task of recognition would be if SRS software designers

could know that the dictation fragment under consideration came from an abdominal CT scan and was describing the liver. The domain of possible words, phrases, and concepts would be narrowed from the entire corpus of radiology down to a manageable subset. This kind of segmentation and restriction would be possible with a system build around report templates specifically targeted by modality, body region, and organ/function.

In the same way that dictation fragments “tagged” with modality and body region will facilitate computer recognition of their content for translation into text, NLP would be better suited to extract more strictly codified meaning. Thus, numeric codes, words from a restricted vocabulary, or ontological constructs would be available for research, quality assurance, and administrative requirements. Because such NLP can be done using results of dictation in narrative text, radiologists would be freed of the cognitive burden of trying to assign codes, select from restrictive menus, or to use rigidly constrained vocabulary. Once coded, the report could be stored and presented in several formats depending upon the needs and preferences of the reader. Among these formats and “readers,” I would include computer databases queried by analytical processing software. Langlotz has described the potential multiple benefits arising from the combination of narrative dictation, speech recognition, and post processing by NLP.⁵⁶

THE TALKING TEMPLATE CONCEPTUAL DESCRIPTION

I have already described some of the theoretical and practical pitfalls of currently available SRS and emerging computerized menu-driven radiology reporting solutions. I propose a “middle way” that combines the best features of each approach so as to recapture efficiency for radiologists, while at the same time achieving structured format, and consistent content. These are two of the goals articulated in the framework for improved reporting mentioned in the introduction.¹ The third goal of improved reporting, standard language and/or codified results, would be achieved through post processing with NLP without disturbing the essential features that foster communication between radiologists and clinicians. The term that I will use for this technological concept is the

“talking template.” As we have seen above, the problem of look away time is perhaps the most vexing to those designing and using new technology for reporting. The talking template directly addresses the look away problem by providing for all navigation feedback to be delivered through audible cues to the radiologist. At the same time, buttons on the dictation device and/or spoken commands would suffice to control the application interface. Thus, the radiologist would never look away from the images during an entire episode of reporting a case. I believe that this would not only save time but would also obviate problems with cognitive overload resulting from having to switch between complex visual interfaces.

The talking template would preserve the familiar and intuitive aspects of narrative description using an internal “checklist.” Here, the checklist would be predefined and contained in the report template rather than in the radiologist’s memory. Any proposed system should contain facility for users to easily design or modify their own templates. Hopefully, radiology groups would agree on a minimal set of standard templates shared between them. Ideally, a set of core templates would be universally available and widely utilized so that reports from any institution would have the same basic content in the same order. The key to achieving such standardization will be for radiologists themselves to design the templates and for them to know that using them would actually be more efficient than the alternatives.

It is important to note that there is no reason that the predefined sections of the talking template must be identical to the way the report is displayed for referring clinicians. For example, radiologists might prefer to have sections relating to all organs or functions contained in the talking template used to generate a report. For readability, several of these sections could be collapsed into a single paragraph when the document is displayed for clinical review. Alternatively, the entire report could be rendered into narrative free-text form with few, if any, headings. At least two of the currently marketed menu-driven reporting systems already do this by translating menu choices into sentences and paragraphs. Considering that clinicians seem to prefer itemized reports, this may not be necessary. One very interesting idea has to do with display of the impression section. During dictation, keep the traditional ordering whereby

the findings are dictated first and the impression last. However, for display or printing of the report, move the impression to the top reflecting its importance. This is possible with the talking template because each section of the report will be unambiguously identified.

For maximal efficiency, a dictation system using the talking template concept would need to be tightly integrated with order entry in the hospital HIS/RIS, the departmental PACS, and the soft copy workstation. A popular term of art for this kind of integration is "radiologist workflow." This one of the core components of an RSNA led initiative called Integrating the Healthcare Enterprise (IHE).^{57,58} However, it should be possible to implement the talking template concept in settings that do not enjoy fully integrated workflow (i.e., those still using printed requests). Partly integrated digital dictation systems often receive examination order information from the RIS and this would be available for automatically deciding on the relevant template to load. At the lowest level of integration, the radiologist might have to select the template from a list at the start of the dictation and verify that they are working on the correct patient's examination. However, after that, the audible feedback and nonvisual navigation features would enable dictating the entire case without having to look back at the dictation interface.

THE TALKING TEMPLATE FUNCTIONAL REQUIREMENTS

As mentioned in the Introduction, I have no prejudice concerning specific technical details for implementing the talking template concept. However, as is evident from the subsequent discussion that I do have strong beliefs about how a digital dictation system should function in terms of the way the radiologist experiences the interface. The most radical part of the talking template concept is how feedback from the system is rendered back to the user. This should be via audible prompts either played back through the microphone speaker or through a headphone. Furthermore, because it is vital that the user know where they are within the template, these prompts will need to be spoken words or phrases. I would leave open the choice about whether the spoken feedback takes the form

of computer-generated speech or is prerecorded (perhaps in the radiologist's own voice).

Coupled with spoken feedback about which section of the template is active, is some method for the radiologist to move around in the template. One possibility is to provide a set of buttons on the microphone for navigating the template. Currently marketed SRS interfaces partly implement this feature, although one can only move forward through a template displayed on the editing screen. Clever use of two buttons or a single "rocker" switch could allow moving forward or backward by single sections or to the beginning or end of the report. The spoken prompts, naming each section, would be triggered upon entering. The traditional record, rewind, and playback functions would serve to work with the dictated text within a specific section. An alternative method for allowing the user to navigate through the talking template is to enable spoken commands. These might be as simple as; top, bottom, next, and previous. More robust interaction is also possible whereby speaking the name of the section would cause the system to move to it and trigger audio feedback if successful. This feedback need not be the spoken name of the section but rather specific beeps or tones. There might be one tone for indicating that the system understood the command and another for failure to match. Finally, there should be some way to listen to the entire report with spoken prompts for each section followed by the dictated content played back in sequence.

A brief example might be helpful to picture how the talking template would work in actual practice. Let us say that radiologist Johnson is dictating an abdominal pelvic CT scan with no comparison done on Mr. Smith, a 47-year-old man with flank pain and hematuria. For simplicity, I will assume that Dr. Johnson's department has workflow integration and she is confident that it is error-free. Upon activating the image workstation with Mr. Smith's examination, Dr. Johnson would hear "Indication." She would dictate "47-year-old male with flank pain and hematuria" and press the button to advance to the next section. She would then hear "Comparison," would dictate "none," and advance to the next section. Now she would hear "Examination" and would dictate "CT scan of the abdomen and pelvis without contrast using renal stone protocol." Upon advancing to the next section, Dr. Johnson would hear "Right kidney and

ureter” and she would dictate “No stones in the collecting system, pelvis, or ureter.” Subsequent sections might be; left kidney and ureter, bladder, small bowel, colon and appendix, other solid organs, and so on. The final narrative section would be the impression, and this would similarly be indicated by an audible prompt.

To finish our example, let us say that Dr. Johnson’s department uses a universal forced choice system for codifying entire reports. This codification might be achieved during the dictation by being a separate section with its own audible prompt. Alternatively, appending this single code might be deferred to the editing and verification steps and might even be required prior to final signature. In an ideal situation, appending a result code of unexpected and or severe would trigger automatic notification of the relevant clinician via text page or a popup message to their HIS account. The clinician would acknowledge receipt and this information would be sent back through the system so as to record it. At appropriate times, Dr. Johnson could check a list of her reports requiring notification and initiate contact for unacknowledged ones.

It would be useful to have the visual interface mirror the navigation and dictation events going on with the microphone. This would allow radiologists who really want to see the report taking shape to satisfy their curiosity. However, I would counsel against doing this too much as it would reintroduce the look away problem and obviate the efficiency gains. Another design choice would concern whether or not to display the recognized text in real time or to defer this for later. For maximum efficiency, it would be important to allow and encourage batch dictation of a series of reports in one sitting followed by display, editing, and verification as a separate step. I am a firm believer in the concept that medical informatics interfaces must serve the clinical needs of their users and be easily tailored to individual preferences. However, the temptation to look at the visual analog of the audible (talking template) interface might be so strong as to be counterproductive. One solution would be to hide the visible interface by default but allow it to be activated temporarily. Another way to foster efficient use of the talking template (or any reporting interface) is to conduct rather more extensive and intensive training of radiologists during initial introduction.

CONCLUSION

During the next decade, the traditional method of producing radiology reports using dictation transcribed by humans will almost entirely be replaced with some type of computerized system. The two current candidates are SRS and the so-called point-and-click or menu-driven interfaces. The design of these interfaces is merging as SRS acquire more sophisticated report shell capability and speech recognition is incorporated into the menu-driven systems. In this paper, I have sounded a cautionary note about the possible adverse effects of currently conceived digital reporting technology. These include decrease in radiologist’s efficiency and impairment of cognitive functioning by eliminating the ability to dictate while viewing images. The “talking template” concept was introduced as a way to design a reporting interface that would obviate these problems and support the primary goal of efficient, timely, and error-free clinical communication. The essence of the talking template design is for all feedback about navigation through the document to be in the form of audible (spoken) cues with little if any need to “look away” to the reporting interface during primary dictation of the case. This should allow completing entire cases and even batches of studies in a single sitting while deferring editing, correcting, and signing reports to a later time. A corollary effect is to allow radiologists to dictate in familiar narrative text and transfer the task of assigning codes and/or categorical choices to NLP performed on the completed reports.

REFERENCES

1. Siström CL, Langlotz CP: A framework for improved radiology reporting. *J Am Coll Radiol* 2(1):61–67, 2005
2. Wheeler PS, Simborg DW, Gitlin JN: The Johns Hopkins radiology reporting system. *Radiology* 119(2):315–319, 1976
3. Simborg DW, Krajci EJ, Wheeler PS, Gitlin JN, Goldstein KS: Computer-assisted radiology reporting: quality of reports. *Radiology* 125(3):587–589, 1977
4. Wheeler PS, Raymond S: The computer-based cumulative report: Improvement in quality and efficiency. *Radiology* 182(2):355–357, 1992
5. Bluemke DA, Eng J: An automated radiology reporting system that uses HyperCard. *Am J Roentgenol* 160(1):185–187, 1993

6. Mani RL, Jones MD: MSF: a computer-assisted radiologic reporting system. I. Conceptual framework. *Radiology* 108(3):587-596, 1973
7. Mani RL: RAPORT radiology system: results of clinical trials. *Am J Roentgenol* 127(5):811-816, 1976
8. Seltzer RA, Reimer GW, Cooperman LR, Rossiter SB: Computerized radiographic reporting in a community hospital: a consumer's report. *Am J Roentgenol* 128(5):825-829, 1977
9. Simon M, Leeming BW, Bleich HL, et al.: Computerized radiology reporting using coded language. *Radiology* 113(2):343-349, 1974
10. Leeming BW, Simon M, Jackson JD, Horowitz GL, Bleich HL: Advances in radiologic reporting with Computerized Language Information Processing (CLIP). *Radiology* 133(2):349-353, 1979
11. Leeming BW, Porter D, Jackson JD, Bleich HL, Simon M: Computerized radiologic reporting with voice data-entry. *Radiology* 138(3):585-588, 1981
12. Bramble JM, Chang CH, Martin NL: A report-coding system for integration into a digital radiology department. *Am J Roentgenol* 152(5):1109-1112, 1989
13. Frank MS, Green DW, Sasewich JA, Johnson JA: Integration of a personal computer workstation and radiology information system for obstetric sonography. *Am J Roentgenol* 159(6):1329-1333, 1992
14. Hu CH, Kundel HL, Nodine CF, Krupinski EA, Toto LC: Searching for bone fractures: a comparison with pulmonary nodule search. *Acad Radiol* 1(1):25-32, 1994
15. Kundel HL, Nodine CF, Krupinski EA: Searching for lung nodules. Visual dwell indicates locations of false-positive and false-negative decisions. *Invest Radiol* 24(6):472-478, 1989
16. Kundel HL, Nodine CF, Carmody D: Visual scanning, pattern recognition and decision-making in pulmonary nodule detection. *Invest Radiol* 13(3):175-181, 1978
17. Banks MS: Neuroscience: what you see and hear is what you get. *Curr Biol* 14(6):R236-R238, 2004
18. Buchel C, Price C, Friston K: A multimodal language region in the ventral visual pathway. *Nature* 394(6690):274-277, 1998
19. McGurk H, MacDonald J: Hearing lips and seeing voices. *Nature* 264(5588):746-748, 1976
20. Shams L, Kamitani Y, Shimojo S: Illusions. What you see is what you hear. *Nature* 408(6814):788, 2000
21. Iles J, Walsh V, Richardson A: Visual search performance in dyslexia. *Dyslexia* 6(3):163-177, 2000
22. Slaghuys WL, Lovegrove WJ, Davidson JA: Visual and language processing deficits are concurrent in dyslexia. *Cortex* 29(4):601-615, 1993
23. Slaghuys WL, Twell AJ, Kingston KR: Visual and language processing disorders are concurrent in dyslexia and continue into adulthood. *Cortex* 32(3):413-438, 1996
24. Ash JS, Berg M, Coiera E: Some unintended consequences of information technology in health care: the nature of patient care information system-related errors. *J Am Med Inform Assoc* 11(2):104-112, 2004
25. Garrod S: How groups co-ordinate their concepts and terminology: implications for medical informatics. *Methods Inf Med* 37(4-5):471-476, 1998
26. Cimino JJ, Patel VL, Kushniruk AW: Studying the human-computer-terminology interface. *J Am Med Inform Assoc* 8(2):163-173, 2001
27. Patel VL, Kaufman DR: Medical informatics and the science of cognition. *J Am Med Inform Assoc* 5(6):493-502, 1998
28. Patel VL, Kushniruk AW: Understanding, navigating and communicating knowledge: issues and challenges. *Methods Inf Med* 37(4-5):460-470, 1998
29. Berg M: Practices of reading and writing: the constitutive role of the patient record in medical work. *Sociol Health Illn* 18:499-524, 1996
30. Ramaswamy MR, Chaljub G, Esch O, Fanning DD, vanSonnenberg E: Continuous speech recognition in MR imaging reporting: advantages, disadvantages, and impact. *Am J Roentgenol* 174(3):617-622, 2000
31. Rosenthal DI, Chew FS, Dupuy DE, et al.: Computer-based speech recognition as a replacement for medical transcription. *Am J Roentgenol* 170(1):23-25, 1998
32. Houston JD, Rupp FW: Experience with implementation of a radiology speech recognition system. *J Digit Imaging* 13(3):124-128, 2000
33. Heilman RS: Voice recognition transcription: surely the future but is it ready? [editorial]. *Radiographics* 19(1):2, 1999
34. Gale B, Safriel Y, Lukban A, Kalowitz J, Fleischer J, Gordon D: Radiology report production times: voice recognition vs. transcription. *Radiol Manage* 23(2):18-22, 2001
35. Sistrom CL, Drane WE: Computer voice recognition transcription in radiology: patterns of usage by individual radiologists in clinical practice [abstract]. *Radiology* 213(P):129, 1999
36. Sistrom CL, Honeyman JC, Mancuso A, Quisling RG: Managing predefined templates and macros for a departmental speech recognition system using common software. *J Digit Imaging* 14(3):131-141, 2001
37. Sistrom CL, Honeyman JC: Template and macro management for radiology speech recognition systems. In: Reiner BI, Siegen EL, Weiss DL (eds). *Electronic Reporting In the Digital Medical Enterprise*. Great Falls, VA: Society For Computer Applications In Radiology, 2003, pp 73-82
38. Sistrom CL, Lanier L, Mancuso A: Reporting instruction for radiology residents. *Acad Radiol* 11(1):76-84, 2004
39. Clinger NJ, Hunter TB, Hillman BJ: Radiology reporting: attitudes of referring physicians. *Radiology* 169(3):825-826, 1988
40. Gunderman RB, Ambrosius WT, Cohen M: Radiology reporting in an academic children's hospital: what referring physicians think. *Pediatr Radiol* 30(5):307-314, 2000
41. Johnson AJ, Ying J, Swan JS, Willicam LS, Applegate B, Littenberg B: Improving the quality of radiology reporting: a physician survey to define the target. *J Am Coll Radiol* 1(1):497-505, 2004
42. Lafortune M, Breton G, Boudouin JL: The radiological report: what is useful for the referring physician? *Can Assoc Radiol J* 39(2):140-143, 1988
43. McLoughlin RF, So CB, Gray RR, Brandt R: Radiology reports: how much descriptive detail is enough? *Am J Roentgenol* 165(4):803-806, 1995
44. Naik SS, Hanbidge A, Wilson SR: Radiology reports: examining radiologist and clinician preferences regarding style and content. *Am J Roentgenol* 176(3):591-598, 2001
45. Sistrom CL, Honeyman-Buck JC: Free text versus structured format: information transfer efficiency of radiology reports. *Am J Roentgenol* 2005; In press.

46. Kahn CE Jr, Wang K, Bell DS: Structured entry of radiology reports using World Wide Web technology. *Radiographics* 16(3):683–691, 1996
47. Kahn CE Jr: Self-documenting structured reports using open information standards. *Medinfo* 9(Pt 1):403–407, 1998
48. Kahn CE Jr: Standard Generalized Markup Language for self-defining structured reports. *Int J Med Inf* 53(2–3): 203–211, 1999
49. Kahn CE Jr, de la Cruz NB: Extensible markup language (XML) in health care: integration of structured reporting and decision support. *Proc AMIA Symp* 725–729, 1998
50. Kahn CE Jr: A generalized language for platform-independent structured reporting. *Methods Inf Med* 36(3): 163–171, 1997
51. Kahn CE Jr, Huynh PN: Knowledge representation for platform-independent structured reporting. *Proc AMIA Annu Fall Symp* 478–482, 1996
52. Wang C, Kahn CE Jr: Potential use of extensible markup language for radiology reporting: a tutorial. *Radiographics* 20(1):287–293, 2000
53. Liberman L, Menell JH: Breast imaging reporting and data system (BI-RADS). *Radiol Clin North Am* 40(3):409–430, 2002
54. Lehman CD, Miller L, Rutter CM, Tsu V: Effect of training with the American College of Radiology breast imaging reporting and data system lexicon on mammographic interpretation skills in developing countries. *Acad Radiol* 8(7): 647–650, 2001
55. Berg WA, Campassi C, Langenberg P, Sexton MJ: Breast imaging reporting and data system: inter- and intra-observer variability in feature analysis and final assessment. *Am J Roentgenol* 174(6):1769–1777, 2000
56. Langlotz CP: Automatic structuring of radiology reports: harbinger of a second information revolution in radiology. *Radiology* 224(1):5–7, 2002
57. Channin DS: Integrating the Healthcare Enterprise: a primer. Part 2. Seven brides for seven brothers: the IHE integration profiles. *Radiographics* 21(5):1343–1350, 2001
58. Vegoda P: Introducing the IHE (Integrating the Healthcare Enterprise) concept. *J Healthc Inf Manag* 16(1):22–24, 2002