Cayley, John. "At the End of Literature." *Grammalepsy: Essays on Digital Language Art*. New York: Bloomsbury Academic, 2018. 211–220. *Bloomsbury Collections*. Web. 26 Apr. 2023. <<http://dx.doi.org/10.5040/9781501335792.ch-015>>.

Downloaded from Bloomsbury Collections, www.bloomsburycollections.com, 26 April 2023, 10:45 UTC.

Copyright © John Cayley 2018. Released under a CC BY-NC-ND licence (<https://creativecommons.org/licenses/by-nc-nd/3.0/>). You may share this work for non-commercial purposes only, provided you give attribution to the copyright holder and the publisher, and provide a link to the Creative Commons licence.

At the End of Literature

Literature is made with language. In certain contexts, literature is proposed as the art of language, its highest art. We may contest the range and extent of literature with regard to practices of language as a whole, and we may not agree that the horizons of these practices coincide with the horizons of art that is or may be made from language. Performative, time-based linguistic practices, for example, may not be accepted as unequivocally within the domain of literature, although they are, nonetheless, embraced and appreciated as aesthetically, culturally valuable at the highest levels. At the beginning of this twenty-first century, works designated as poetry or (literary) fiction are generally accepted as literature, all but regardless of quality if not subgenre. Dramatic writing, however, must contest its place as literature to the extent that it is readable in a form that submits to textual practices, effectively print publication—taking this form according to relatively arbitrary conventions of transcription and, indeed, remediation (as “literal” literature)—and also, importantly, in so far as it is studied and critiqued within literature’s discursive frameworks: within the university or the world of letters. If we consider actual artifacts of dramatic writing to be oral performances of language that come to exist, chiefly, in aurality (as much as they do in other media—visual, gestural, architectural, etc.), then it becomes possible to acknowledge that our conceptions of literature and of language art, particularly in terms of the cultural significance of specific artifacts, are not media-agnostic. The relationships of literature with media are historically determined, culturally contingent, prejudiced, and, I will argue, disordered with respect to technological developments. The existing relationships generate aporias that threaten to become critical over time. They deform and distort our appreciation of language art in other media. They cause us to ask ourselves how and why should we value the significance and affect that such work generates.

This problematic—how to appreciate electronic literature or, more generally, language art in digital media—has preoccupied theorists and

practitioners since the beginning of the undoubtedly literary history associated with these practices in the mid-1990s. I say “undoubtedly literary” because it is a matter of record that electronic literature found early, if not unprecedented, theoretical and critical support for its nascent practices in the discursive space of the universities. Although attention to media specificities (distinctly plural) has been properly claimed in order to take into account the incorporation of other media into practices of literary composition and reception, less attention has been directed to any better understanding of the underlying medium of the underlying art. This medium is language and its underlying art is an art of language. The introduction of other media—into practices for the composition and reception of language art—has demonstrably and necessarily broken conventional form. In themselves, such breaks render their artifacts no longer (exclusively, traditionally) literature. Theorists and practitioners of the new forms claim that, in these circumstances, literature must change. But breaking conventional form is nothing new, even across media. What is new—supporting the original claim—is some historically important manner in which literature is called to change—paradigmatically, conceptually, fundamentally. Literature needs to become “electronic,” by which we mean (with hindsight) that it must come to terms with the digitalization of everything.

The digital—inevitably misrepresented as “electronic” for the rhetorical purposes of the claim that literature must change—is not a medium. More precisely, it is not a medium of interest to the majority of theorists or practitioners of those arts for which language is the medium. There are aesthetic practices of computation and of properly digital art, with respect to which the digital can be accounted as *a* or *the* medium, but only certain specialist practices of electronic literature incorporate computational aesthetics significantly or affectively. For media taken as the plural of medium, the digital is, rather, a prevalent and privileged framework and network for any and all media. These media—color, shape, texture, sound, and so on—are encoded in sometimes complex, structured binary transcriptions that render these digitized representations accessible to and manipulable by computational, digital affordances. For media such as these, which are referred to substantive material, digital representations are problematic in many interesting ways characterized by our understanding of significant and affective differences between analog and digital objects or artifacts, and yet there is a phenomenological coherence in terms of the human experiences of these things across the analog–digital divide since they are also necessarily referred to human perception—of color, shape, texture, sound, and so on.

The relationship between linguistic artifacts and digitization is, on the other hand, singular. A string of bytes that represents a color, however structured by coding conventions, is not the color itself. By contrast, a string of bytes representing a string of letters and punctuation is language, ontologically, in so far as it is humanly readable. There is no essential

difference between any instance of language as it is embodied “here” and “now” on the page or surface in front of you and how it is encoded as a string of bits inside your machine. Its existence as language is entirely dependent on your ability to read it. If you are able to read traces (grammē) of this language on any other “surfaces” within any part of a computational system, your reading brings that language into being. The string of bits digitized from existing systems of inscription is always already structured as traces of language that are, in principle, if not in practice, readable. It is not encoded so as to enable the rendering of an object in another medium that *because of this rendering* becomes perceptible as an instance of an object in that medium. Perception of digitization is not perception of what it encodes. The digital representation of color and sound, for example, is not perceptible as such within digital systems. By contrast, when presented with traces of language, in any material form, all the human subject needs to do is read. Any perceptibility (or not) with regard to the material form in question is irrelevant except in so far as it simply enables or disables actual reading.

What is at stake for language and language art in digital media is not a supposed ontological distinction between language and digital language in so far as this is a function of digitization. Whereas it does make sense to speak of a distinction between yellows (yellows in our perceptible world) and digital yellows (encodable and renderable yellows), it makes no sense to speak of a purported distinction between words and digital words. Instead we must turn our attention to effects of digitalization on the substantive media that can support traces of language and their potential for human reading.

Within the much wider domain of linguistic practice, what has occurred, indisputably, since the post–Second World War rise of distributed computation is, fundamentally, the digitization of typography and typographic design, the digitization of *particular aspects of visibility* that are structured so as to support linguistic practices that derive, for the most part, from print-based textuality. In general, and historically, when we speak of electronic literature we speak of a textuality that has activated certain digital affordances with respect to digitized typography. In print, typographic visibility is static, fixed, although it may be spatialized in a number of ways so as to influence or inflect reading practices and strategies. By contrast, even with relatively basic peripherals, digitized typography has nearly all the affordances of print and is provisioned, additionally, with a wide array of dynamic potentialities. Text in digital media can move and change. It’s as simple as that. It is important, however, to recognize that this is not a difference in *what* is or can be read but in *how* and *when* it is read. The digitization of typography has given us new expressive structures for temporalities that have the potential to influence and change the fundamental events of language: our events of reading.

Thus, the fact that there is no ontological distinction between language and digital language does not mean that digital textuality—digitized typography—as compositional media and expressive form can be reduced

to textuality as modeled by the more constrained expressive potentialities of print. That ship has long sailed. And if literature is a practice that is determined, chiefly, by material cultural formations that orbit practices and conventions of reading, then it is literature that faces its ontological challenge with respect to digitalization. Electronic literature is, precisely, no longer literature; if it is anything, then it is *digital language art*, although currently it still struggles within the gravity of an “electronic literature” that is overdetermined by aesthetically motivated language expressed in the substantive medium of digitized typographic visuality. Even as such, within the constraints of existing practices, the digitization of typographic visuality tends to facilitate new ways of reading, especially less familiar temporalities of reading, and new relationships between reader action and what is read (hypertextual and conditional linking). For most readers, even including critics, literary scholars, and digital humanists, these strategies trouble existing traditions of literary reading without yet insisting that literature itself be called, seriously, into question.

Throughout this thinking, a particular conception of reading is crucial. I speak of reading in a specific technical sense. I use *to read* and its cognates—in a manner fully consistent with its etymology—to refer to whatever it is that we language animals do when we discern and interpret linguistic forms, *regardless of support medium*. This is not the type of metaphoric usage that obtains when we speak of “reading” a painting or a dance. It refers to the process of grasping and understanding traces of language as such in any medium. In this thinking, once it comes into existence, language is not only discrete and articulated, it is distinctly separable from other phenomena of the perceptible world, made and marked by what Jacques Derrida indicated as *différance*.¹ Virtual linguistic forms establish a break with the perceptible matter of which they are formed precisely in that catastrophic, no-turning-back moment when they are grasped as language by both the language animal who makes the traces and a language animal who reads them. I call this process grammalepsis and I consider it to be generative of language, ontologically. Reading brings language into substantive being as instances of interhuman potentialities. To clearly distinguish reading in this sense from the subsumed and more specific activity that we undertake when, typically, we visually scan and interpret instances of writing, we could use the phrase *grammaleptic reading*, but so long as we recall, throughout this thinking, that this special sense of reading is equally what we do when we hear and understand spoken language in aurality (or, for example, when deaf communities read sign language or blind communities braille), I may use “reading” on its own, with the inevitability of grammalepsis comprehended.

Once we are able to accept (grammaleptic) reading in this sense as constitutive if not ontologically generative with regard to language, this is when it becomes possible to appreciate more fully certain potentialities of digitalization, certain anticipated effects on language and its arts at this

particular historical moment. We have argued that digitization changes our modes of relationship—transaction and interaction—with the support medium for language rather than with language itself, *how* and *when* we read rather than *what* we read. What, then, happens when there is, in the domain of digitalization, a catastrophic (no-turning-back) *convergence of readabilities in terms of grammalepsis* with regard to the two distinct, if imbricated and culturally implicated, media that support language: visuality and aurality?

At this point we must pause to consider certain relationships between language, its support media, and the language animals that bring language into being—ourselves. Language is something that, to the present extent of our knowledge, only humans *have*.² Our species has language. It evolved to have language in a manner that is still imperfectly understood, although there are particular characteristics of this evolved condition that can be specified. There are distinct, implicated morphological traits that we have and other animals do not. From my reading, I take the most significant of these, apart from larger brains (which may not be as crucial as we suppose), to be: a double-articulated oral cavity and larynx, and a spinal column with a significantly greater diameter. In concert with *Homo sapiens*' larger brain size (and perhaps many other factors), these traits allowed us to *have language* because we were suddenly, in terms of evolutionary time, able to make a sufficient number of distinct vocal sounds—sufficient for vocabulary and grammar commensurate with language as we know it—and because a larger spinal column allowed nerve cells and interneurons to establish the fine control over our lungs that was also required for articulation.³ This happened to our species relatively recently in evolutionary time. Effectively, we have had the potential for language baked into us very, very recently and there are unlikely to be any foreseeable genetic changes in our species that will significantly alter our disposition with respect to language. The point being that we are genetically predisposed to *have language* as a function of traits that operate *in aurality*. If we have adopted visuality as the support medium for particular linguistic practices of what we call writing, this is merely learned, a function of civilization.⁴

It is well known and much discussed that Plato considered writing to be a pharmakon, poisonous to practices of language—particularly language as humanly embodied praxis and cultural memory. And yet, in its other aspect writing-as-pharmakon was rendered therapeutic by civilization. This is, of course, a grand narrative, played out in philosophy following and reading Plato, most particularly in the thought of Jacques Derrida and Bernard Stiegler.⁵ Writing and, subsequently, literature as linguistic practice in support of civilization were rendered therapeutic precisely because they restructured the temporalities of language as well as enabling the potentialities of index (random access facilitated by sublexical orders giving more or less instant access to significant and affective textual material) and archive. Clearly, writing allows virtual linguistic performance to survive—in

temporal extension—not only the actual performance of its makers but also the memories of particular individuals who have *read* (grammaleptically) particular linguistic performances. This temporal affordance—hypostatic memory or hypomnesia coupled with index and archive; preserving and conserving both language itself and these other two features—allowed writing, ironically, to predominate as *the* privileged *literal* index of logocentric presence and authority: history, philosophy, civilization.

Putting it far too plainly: as the course of human history and culture proceeded, language in aurality was not able to participate as effectively as writing—as language in persistent visibility—for the constitution and maintenance of civil and imperial institutions. Until, that is, just about now, at this time of writing, in the 2010s. This decade has witnessed the advent of *transactive synthetic language* in aurality.⁶ Contemporary computation has finally achieved robust voice recognition and acceptable speech synthesis, all implemented over network services having access to vast corpora of natural linguistic material with NLP affordances. Historically, I argue, this is a turning point for our—the language animal’s—practice of language in the world, since, for one thing, this world now also contains, crucially, humanoid language and new entities that perform, consume, and transact with both language as such and humanoid language.

There might arise a certain objection to my dating of the proposed paradigmatic shift, in that synthetic (computed) language has played a part in the history of computation since its beginnings, including, foundationally, in the exemplary abstracted scene of *writing* that is the Turing Test, for which the *withholding* (by, at the time, teletype) of any embodied voice is crucial *for the test*, since a voice and body would simply give the game away.⁷ In a sense, the advent of systems that we humans agree are able to recognize our voices and respond with—gendered and identifiable—voices of their own forecloses the Turing Test and marks it as having already been passed within the duration of any acceptable initial transaction. It is the system’s voice—recognizing and producing virtual language and doing so *necessarily* instantiated in aurality—that is sufficient to establish for us human animals that the system is specifically embodied as, at least, humanoid, and certainly as having (or seeming to have) something that only humans have. The historical moment for our new relationship with language had to wait for this milestone of humanoid embodiment, in and as the voice of articulated aurality, perhaps also as the evolved return and reincarnation of a repressed aurality. And for the electronic *literature*, that we have troubled and recast as digital language art, this turning point requires us—practitioners and scholars—to better understand what it is that “the digital” has done for language. It has not (yet), as we said before, established an ontologically distinct (digital) language as such; rather, it has reconfigured the relationship between language and its preferred substantive media of support. More than this, it suggests that we rethink, and shift our attention to the *other*

culturally predominate substantive support medium for language. The digital now, historically, forces us to rediscover the voice as articulated aurality in an artifactual and programmable configuration that, in computationally implemented principle, is every bit as manipulable and extensible, as subject to index and archive, as capable of temporal restructuring, as is writing. Transactive synthetic language is a whole new scene for the art of language in general, and for digital language art specifically.

As we begin to shift our attention from theory of language and media toward new practices of language art, it is important more closely to consider what it is that I claim is happening with regard to language in aurality as it is grammatised—subjected to algorithmically implemented processes of grammalepsis—by contemporary computation.

Language has a singular relationship with its substantive media of support. For V. N. Vološinov and certain of his followers, there is such a thing as “semiotic material” and any sign—in my own terms *anything* that has been read grammaleptically—becomes a token of this semiotic material.⁸ Natural languages are socialized, agreed, enculturated systems that are entirely composed of “semiotic material” in this sense. There is a constrained permeability of substantive things that may be on their way to becoming signs, becoming, that is, actual semiotic material. The signs and tokens of natural languages are, on the other hand, always already signs for the language animals that encounter and interact with them, achieving this in a social context that necessarily involves other language animals. Compare a particular gesture of the hand, say. A gesture may already be a sign—it might be conventionally understood in a particular culture or it might be (always already) a sign in a natural (sign) language—but a gesture may also be on its way to becoming a sign, something we don’t “get,” something that needs more work and practice, to get right, to be able to express, significantly and affectively, whatever it hoped to express. It fails in this until it is grasped, until it succeeds. It fails until it is read grammaleptically.

The written forms of any natural language have long ago passed beyond this underlying scene of semiotic trauma and socialization, to the extent that the chains of tokens of language-in-visibility (strings) enter into the domain of purely formal semiosis—computation—in a wide variety of processes that are, fundamentally, “lossless” ontologically. If you can write it, then you can encode it. And, as we showed earlier, the language-as-visual-graphemes—on paper, on screen—is ontologically identical with any language-as-digital-encoding that underlies it. In either case, what makes the language exist *as such* is its potential to be read, grammaleptically, by language animals.

If we are repeating ourselves and somewhat belaboring these concepts, this is due to the necessity to distinguish—in the domain of aurality—between the digitization of sound and the digitization of language-as-aurality. There is a significant critical literature devoted to the media archaeology of recorded sound, and this is often seen in terms of a refiguration of digital audio

recording and transcription. In this literature, there is clear understanding and analysis of distinctions between analog and digital recordings, with important implications to be drawn. Nonetheless, once these have been elaborated, there may be a misdirected tendency to believe that because the digitization of sound encompasses and comprehends the digitization of linguistic sound, then it has comprehended the digitization of language-as-aurality. But this is not the case.

As set out above, the digitization of sound is constituted by encoding the forms of a substantive medium which then require to be rendered before they can be appreciated as such, as structured sound. You cannot hear the encoded version. Digitized linguistic sound is no different. The encoded version cannot be heard, much less read, grammaleptically, as language. Any grammalepsis of digitally encoded linguistic sound can only occur during a separate, subsequent process, *after* it is rendered into the world as sound. At this point, the sound *may* be read and understood, by language animals, as language-in-aurality.

In our present historical moment, the 2010s, robust automatic voice recognition is fast gaining currency in the digitalized world, currently to be qualified as, chiefly, the global Anglophone world. It is this facility—automatic voice recognition—that enables the actual digitization of language-as-aurality. Evidence that this facility was beginning to be operational dates back to early attempts at automatic dictation/transcription systems, voice command interpretation for personal computers, and, especially, automated voice-activated telephone answering systems. On mobile devices, Apple's Siri was a breakthrough but, for our purposes, as research-based practitioners and theorists, it is the Amazon Echo and its Alexa Voice Services (AVS) that provide the first widespread, operational, free-standing, networked, and programmable infrastructure, allowing us to understand, practically, the effects and potentialities of digitized language-as-aurality.

As a point of operational fact, the Amazon Echo and AVS enact precisely the two-stage process of digitization for language-as-aurality that we alluded to above. Not only does this configuration of the AVS infrastructure demonstrate that the procedures are distinct, it also signals our always insufficiently acknowledged reliance on network services, with the asymmetric balance of agential power and centrality that this implies. The two procedures are separately *located*. I speak to an Amazon Echo. The device, locally and in “real time,” optimizes its array of microphones to capture as digital audio a segment of—purportedly—linguistic sound that was prefaced by one of its (currently three) “wake” or trigger words. These wake words are the only fragments of sound that the device itself, locally, is able to read grammaleptically as semiosis: a command to record, until a space of relative silence is encountered. Within the device this digitized audio is encoded as an optimized MP3 file, and it is this digital *audio* data that is transmitted over the network to the cloud-based services of AVS. “In the cloud” this digitized linguistic sound is “recognized,”

which is to say tokenized by automatic grammalepsis into, currently, word-sized, serviceable “atoms” of machine-modeled natural language. The details of this process are proprietary, although many aspects of the underlying research could be set out and exposed. The pragmatic approach implied above by “word-sized” is an educated guess. What we know as a certainty, because the AVS cloud services supply (they “return”) transcriptions of what the system “heard” grammaleptically as *text*, as potentially readable language that is materially identical with all the digitized writing that constitutes the most significant material of networked digital culture: the documentary internet as we know it.

It is important to acknowledge that this service—which we are proposing as, potentially, of momentous, paradigm-shifting cultural efficacy—does not deliver understanding. This is not the hermeneutically enhanced grammalepsis of reading as it is performed by fully enculturated language animals. In the theory and practice of automatic voice recognition, this is deferred, researched, and explored as “automatic understanding,” more firmly in the speculative camp of machine learning and artificial intelligence. Automatic speech recognition does, however, achieve the digitization of language-as-aurality, which means: language animals may perform in a manner toward which they are genetically disposed and what they say is, in principle if not yet perfectly, automatically recast in an encoded form, subject to digital affordances, that is materially identical to text, to writing, to all the strings of language that are now humanly readable in the realm of computation and our increasingly predominant digitalized culture. I am tempted, provocatively, to say (to write!) that socialized automatic speech recognition transforms human linguistic performance into literature. Except that I imagine that such practices, for aesthetic, significant, and affective purposes might one day have no human need for literature as such. Its greatest work will always already—and would not Shakespeare scholars agree with me here?—be aurature.

Before the 1990s, language-as-graphemic-or-typographic-visibility was already digital. (Since the very advent of writing systems, I, for one, would argue.) Algorithm and formal procedures of many kinds had long been applied to natural human language in this form, as writing, as literature, including and particularly for aesthetic effect. It was the enculturation of widespread media-agnostic digital affordances that, in the 1990s, allowed specialist practitioners and scholars to characterize what were essentially quantitatively and peripherally rather than qualitatively distinct reconfigurations of literary material as, speculatively, “electronic literature.” Digital affordances allowed practitioners and scholars to do new things with old words, to an extent that rendered some of these new things interesting and exciting. But reading as such did not change. Nor will it fundamentally change until the language animal that is definitive of reading has time to evolve. What did change, even in the 1990s, was the configuration of the scene for linguistic poesis—

the hows and whens of reading and writing. This was and is momentous enough, but hearing and speaking go on much as they have done, and the predominance and momentum of reading and writing traditions were and are minimally deflected. Even now, the most industry and energy that has been expended on the remediation of literary practices has been applied to artifacts that support the tradition of the book, of print-based, typographic media—those emulators, images and mirrors of typographic artifacts that, in English, go by the disfigured name of “ebooks.” Ebooks are with us, for the time being and foreseeable future, but at the time of writing growth in their popularity and dissemination has slowed. Over roughly the same period there has been significant growth in the reading of audiobooks despite the fact that culture predisposes these readers to an anxiety concerning whether or not they have actually read what they are reading.⁹ As of 2018, the audiobook is not digitally inscribed as language-in-aurality. It is, rather, digitized audio with minimal digitally manipulable articulation corresponding most commonly to the punctuation of books at the level of the chapter or subtitle. Nonetheless, the reading of audiobooks represents a measurable shift in the culture of reading as a whole, and this development coincides with what I speculate will become the socialization of automatic speech recognition such that the aurality of existing books is or will be grammatized at the level of (at least) the word, and—to indicate merely practices that are already available to certain readers—speech synthesizers are or will be able to present this language-as-aurality to human readers directly, automatically. We will have the option of reading in this newly articulated aurality.

If we can read in aurality then, as language animals and language artists, we can compose in aurality. We can begin to make an aurature that is formally, philosophically, ontologically identical with the literature we have inherited, an aurature that will reconfigure and redefine the archive without in any way sacrificing readability in general or the specific mode of readability that has been established by literacy. The full civilizing potential of this prospect—an aurature embodying facilities with language that are attuned to our genetic disposition as language animals—is available to us only due to crucial developments in digital culture and contemporary computation. Hence, we can affirm that practices of *digital language art*—especially in the reconfigured support media for language as an aesthetic medium—at least makes sense, and may also imply, I believe, cultural and social imperatives. Practitioners and theorists must learn and grasp those computational affordances that will allow them, fully, to participate in, to guide, and to enhance cultural and social developments that will otherwise proceed without their contributions, and risk downplaying aesthetic practice at the expense of what are supposed to be more substantive and instrumentally secure benefits. What we do not want is to remain the electronically literate writers of a history in which we find ourselves at the end of all literature, with no viable media for the art of language.