

Machine Reading in the Mexican Colonial Archive

OCR and the Primeros Libros

Hannah Alpert-Abrams, University of Texas at Austin

Transcription

“The allure of the archives passes through this **slow and unrewarding artisanal task of recopying texts**, section after section, without changing the format, the grammar, or even the punctuation. Without giving it too much thought. Thinking about it constantly. As if the hand, through this task, could make it possible for the mind to be simultaneously an accomplice and a stranger to this past time and to these men and women describing their experiences. As if the hand, by reproducing written syllables, archaic words, and syntax of a century long past, could insert itself into that time more boldly than thoughtful notes ever could. Note taking, after all, necessarily implies prior decisions about what is important, and what is archival surplus to be left aside. The task of recopying, by contrast, comes to feel so essential that it is indistinguishable from the rest of the work. An archival document recopied by hand onto a blank page is **a fragment of a past time that you have succeeded in taming**. Later, you will draw out themes and formulate interpretations. Recopying is time-consuming, it cramps your shoulder and stiffens your neck. But **it is through this action that meaning is discovered**.

Arlette Farge, *Le Goût de l'archive*

How is automatic transcription a process of:

Labor
Textual Taming
Interpretation
?

Case Study: Primeros Libros

Los Primeros Libros de las Américas

Impresos Americanos del siglo XVI en las Bibliotecas del Mundo



Ocular: Historical Document Recognition System

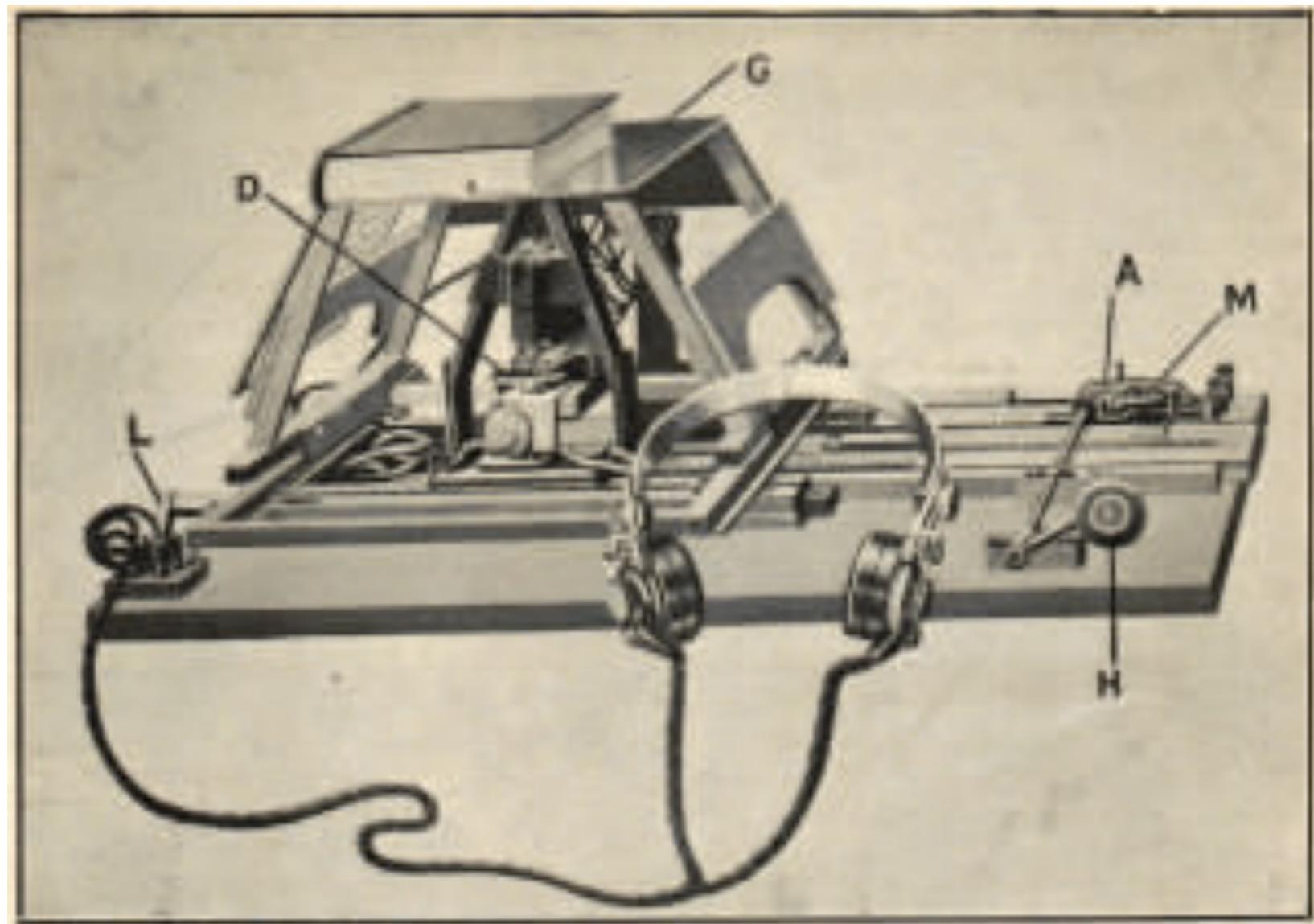
Taylor Berg-Kirkpatrick et al., UC Berkeley (2013)

1. Font Model
2. Language Model

Labor

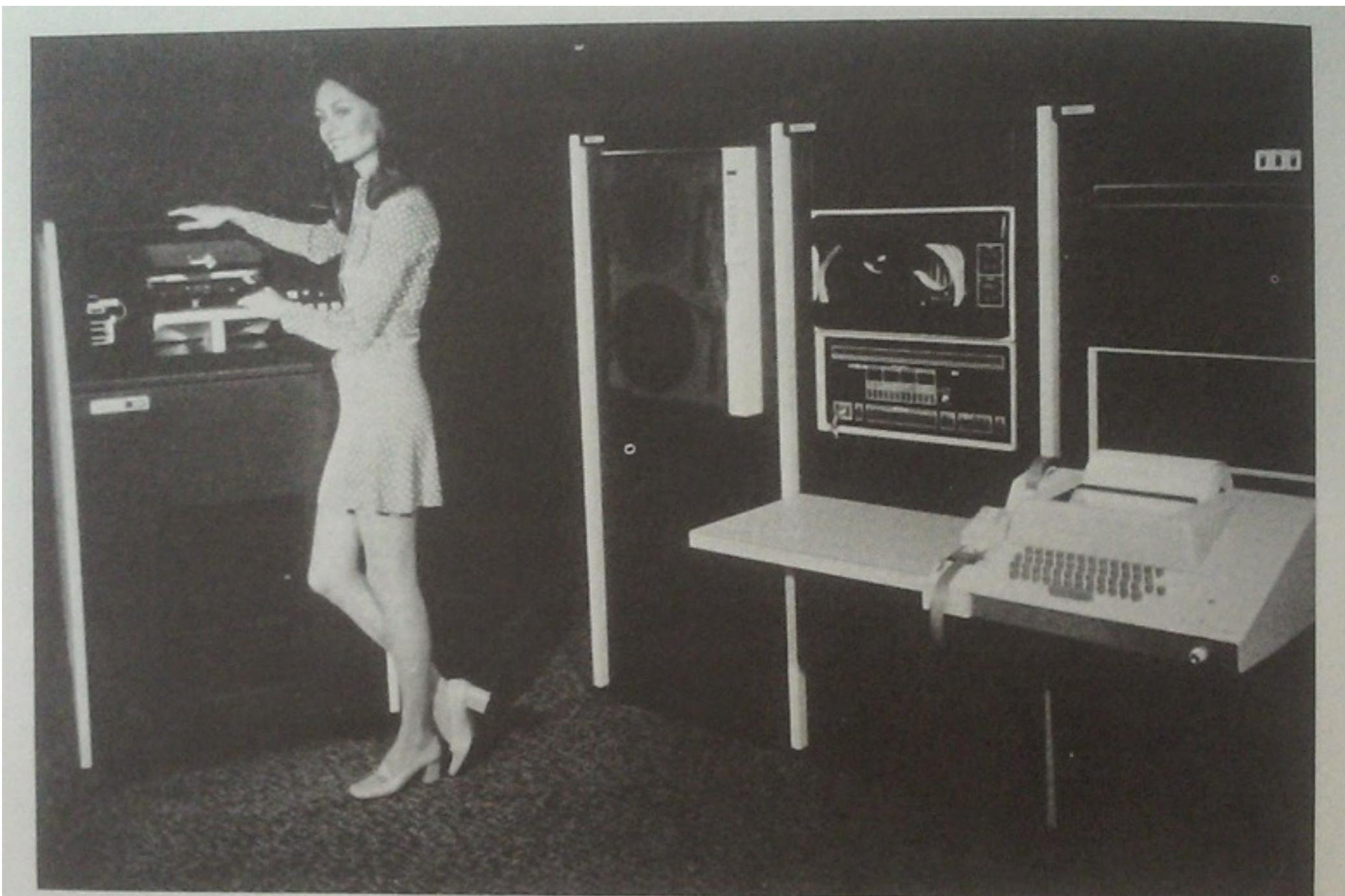
“this slow and unrewarding artisanal task of
recopying texts”

Labor



Schematic of optophone from *Vetenskapen och livet* (1922)
[Mara Mills, "Optophones and Musical Print"]

Labor



Scan-Data 2250, friction feed, flying spot scanner, feature Mixed Media On-line to Scanplex. (1975)

Schantz, *The History of Optical Character Recognition*

Labor

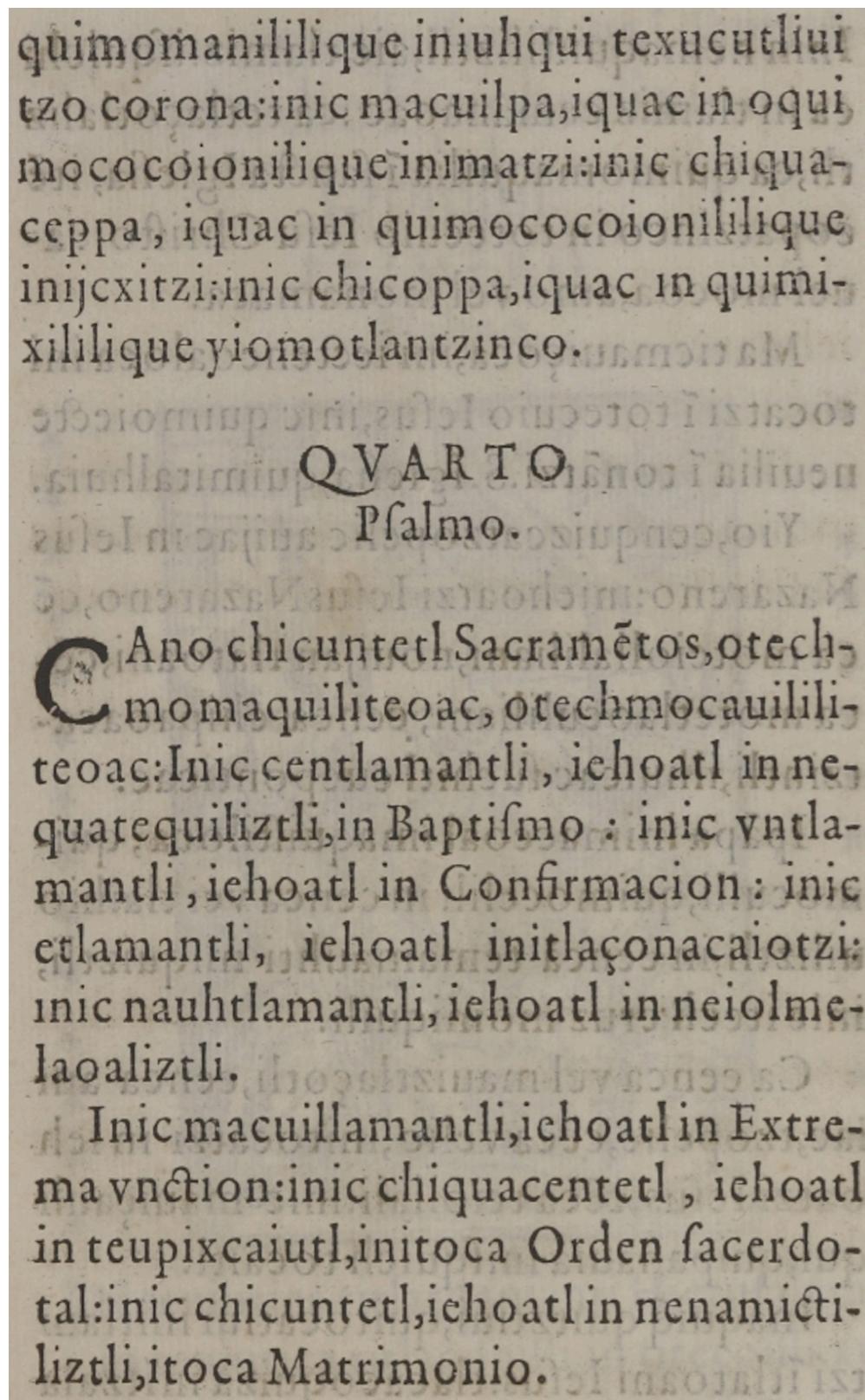
The action of automatic transcription is located not in the reader but in the machine.

This labor is in the service not of the individual citizen, but rather of the corporation or the institution.

Textual Taming

“An archival document recopied by hand onto a blank page is a fragment of a past time that you have succeeded in taming.”

Textual Taming



glutmom an ill liquefiniti liquift executlini
exorcoronawinic magnilpa, square in equi-
me cocotom liquefinimatzionic gluqua-
ge PPa., square in quimoco co fourth lique,
in flexitz is interchicop Pa., square in quirmi-
xist liquefy somedants inco-

Q. VAILTOs.

If salmo-

ares. An orchic unred Sacrameters were else
35.2 momaquiliteo are were climrocaruisili-
teorage Intergentilamandis, tie board isn't new
quate quilted it in Raptisme's snig Vada-
m and swig broad in Confirmacion's inter
girlamandis wife broad. Trinidagon arcators is
internault dramandi, is board in petrol me-
ss oasis di-

stries magnilla man divierboard in Extte-
marvnctions interchiquage needs, tehead
in teup is casud, introga Orders facetts-
tasonic chic unred wire broad in netta micri-
b 450.45 454 Mattonomo-

Textual Taming

	mentira	méjica
Without orthographic variation handling	mentira	merita
With orthographic variation handling	mentira	m̄tira

Textual Taming

Ocular “tames” historical documents by:

Preferring ultra-diplomatic data

Codifying a statistical model of orthographic variation

Interpretation

Recopying is time-consuming, it cramps your shoulder and stiffens your neck. But it is through this action that meaning is discovered.

Interpretation

los confessores pregunten a los Mercaderes o Indios ricos, si han dado a logro: para mandar les restituyr lo que asi vuieren llevado. A y proprio vocablo de logro, que es, tetech-tlaixtlapanaliztli, tetechtlamieccaquixtiliztli, y para dezir diste a logro? Cuix tetech otitlaixtlapan, cuix tech otitlamieccaquixti?

¶ El confessor pregunte siempre a su penitente si cumplio la penitencia que le fue impuesta en la confession passada, y conforme a lo que a lo que le pareciere conuenir asi le mande q la cumpla, o se la comute. Por que el confessor segundo puede comutar la penitencia impuesta por el primer confessor, sin oyr los pecados: por que piadosamente se puede creer y presumir, que el confessor primero no disぐstarà. Y mucho mejor y mas segura cosa es comutarla en alguna indulgencia, de las que se ganan haciendo lo que contiene la Bulla [si el penitente la tuuiere] asi lo dice Nauarro in Consilijs lib.5, Confilio. 21. y Henrico Henriquez, Tomo.1. de Penitentię Sacramēto, lib. 5. cap. vltimo. Dōde dize, Parochus seu equa lis confessariis ea ratione potest facilius comutare: quia prior confessarius censetur interpretatiue concedere eam licentiam, aut ex debito.

los confessores pregúnten a los Mercaderes o Indios ricos, |i han dado a logro: Para mandar les restituyr lo que aí víueren llevado.

A y proprio vocablo de logró, que es . tetech-tlaixtlapanaliztli, tetechtla miec caquixtiliztli, y para decir de te a logro? Cuix tetech otitlaixtlapan, cuix tech otitlani teccaquixti?

8. El confeſſos pregunté |iempre a |u penitēte |i cumplió la penitencia que le fue impuesta en la con feſſión pa||ada, y conforme a lo que á lo que le pareciere conuenir aí le mande q la cumpla, o |e la comute. Por que el confeſſor legundo puede comutar la penitencia impuesta por el primer confeſſor, |in oye los pecados; por que piado|amēte |e puede creer y pre|umir, que él confeſſor primero no dis gu|tárā? Y mucho mejor y más |egura co|a es en matarla en alguna indulgencia, de las que |e ganan haviendo lo que contiene la Bulla ē |í el Penitēte la tuuiere; i afi lo dize Nauarró in 1190 filiis lib. 7. Conſilio. xi. y |í en rico Hén - 149004, Tonto, i de Penitētio Sacramēto, lib.

7. "AP. X." 11049. Dōde dize, Parochus |eu e qua 117 "9016114rtus ea ratione poteſt facilius co - 191174164 ju74 prior confeſſarius cen|etur in - Æ" "P10140116 concedere eam licentiam, aut ex I. I.) ¿88

Interpretation

In Dios, cá Tettatzin, Tepiltzin, Spiritu Jan-

cto, in ieixtin personas çan huel iceltzin Dios

In **Dios**, cá Tettatzin, Tepiltzin, **Spiritu |an-**
cto, in **ieixtin per|onas** çan huel iceltzin **Dios**

Spanish

Latin

Nahuatl

Interpretation

el misterio de la sanctissima Trinidad, creyendo que, meteihittotica. q. d. Trino, y que la sobre dicha proposicion conforme a esto, q. d. no ay mas de vn Dios que es Padre, Hijo, y Spiritu sancto trino en personas, y uno en

el mi|terio de la |ancti|sima Trinidad, creyendo que, ineteihttotica . q' . d . Trino, y que la |obre dicha propo|sición conforme á e|to, q. d. no ay más de vn Dios que es Padre, Hijo, y Spiritu |ancto trino en per|onas, y uno en

Spanish
Latin
Nahuatl

Interpretation

James Lockhart
Louise Burkhart

“Loan words” and language identification are marks of cultural contact, assimilation, and interpretation.

Conclusion: OCR as Digital Documentary Edition

History of OCR

Arlette Farge:

Labor
Textual Taming
Interpretation

The false transparency of computing
- Wendy Chun, *Programmed Visions*

Digital Documentary Editions

Leaves of Grass.

I CELEBRATE myself,
And what I assume you shall assume,
For every atom belonging to me as good belongs to you.

I loaf and invite my soul,
I lean and loaf at my ease . . . observing a spear of summer grass.

Houses and rooms are full of perfumes . . . the shelves are crowded with perfumes,
I breathe the fragrance myself, and know it and like it,
The distillation would intoxicate me also, but I shall not let it.

The atmosphere is not a perfume . . . it has no taste of the distillation . . . it is
odorless,

It is for my mouth forever . . . I am in love with it,
I will go to the bank by the wood and become undisguised and naked,
I am mad for it to be in contact with me.

The smoke of my own breath,
Echoes, ripples, and buzzed whispers . . . loveroot, silkthread, crotch and vine,
My respiration and inspiration . . . the beating of my heart . . . the passing of blood
and air through my lungs,
The sniff of green leaves and dry leaves, and of the shore and darkcolored sea-
rocks, and of hay in the barn,
The sound of the belched words of my voice . . . words loosed to the eddies of
the wind,
A few light kisses . . . a few embraces . . . a reaching around of arms,
The play of shine and shade on the trees as the supple boughs wag,
The delight alone or in the rush of the streets, or along the fields and hillsides,
The feeling of health . . . the full-noon trill . . . the song of me rising from bed
and meeting the sun.

Leaves of Grass.

I CELEBRATE myself,
And what I assume you shall assume,
For every atom belonging to me as good belongs to you.

I loaf and invite my soul,
I lean and loaf at my ease . . . observing a spear of summer grass.

Houses and rooms are full of perfumes . . . the shelves are crowded with perfumes,
I breathe the fragrance myself, and know it and like it,
The distillation would intoxicate me also, but I shall not let it.

The atmosphere is not a perfume . . . it has no taste of the distillation . . . it is
odorless,

It is for my mouth forever . . . I am in love with it,
I will go to the bank by the wood and become undisguised and naked,
I am mad for it to be in contact with me.

The smoke of my own breath,
Echoes, ripples, and buzzed whispers . . . loveroot, silkthread, crotch and vine,
My respiration and inspiration . . . the beating of my heart . . . the passing of blood
and air through my lungs,
The sniff of green leaves and dry leaves, and of the shore and darkcolored sea-
rocks, and of hay in the barn,
The sound of the belched words of my voice . . . words loosed to the eddies of
the wind,
A few light kisses . . . a few embraces . . . a reaching around of arms,
The play of shine and shade on the trees as the supple boughs wag,
The delight alone or in the rush of the streets, or along the fields and hillsides,
The feeling of health . . . the full-noon trill . . . the song of me rising from bed
and meeting the sun.

Digital Documentary Editions

ON

Loft.

Tta he would „i„e f for i f,, , .TM •

The facred Troth, to Fabl e and old & J

WI he perplex'd the things he would expfain^ *

J

Work

£j asd ,f q<><> always what is weJJ,

And by ill imitating would excell!)

Might hence perfume the whole Creations dav

To change m Scenes, and /how it in a Pl av 7

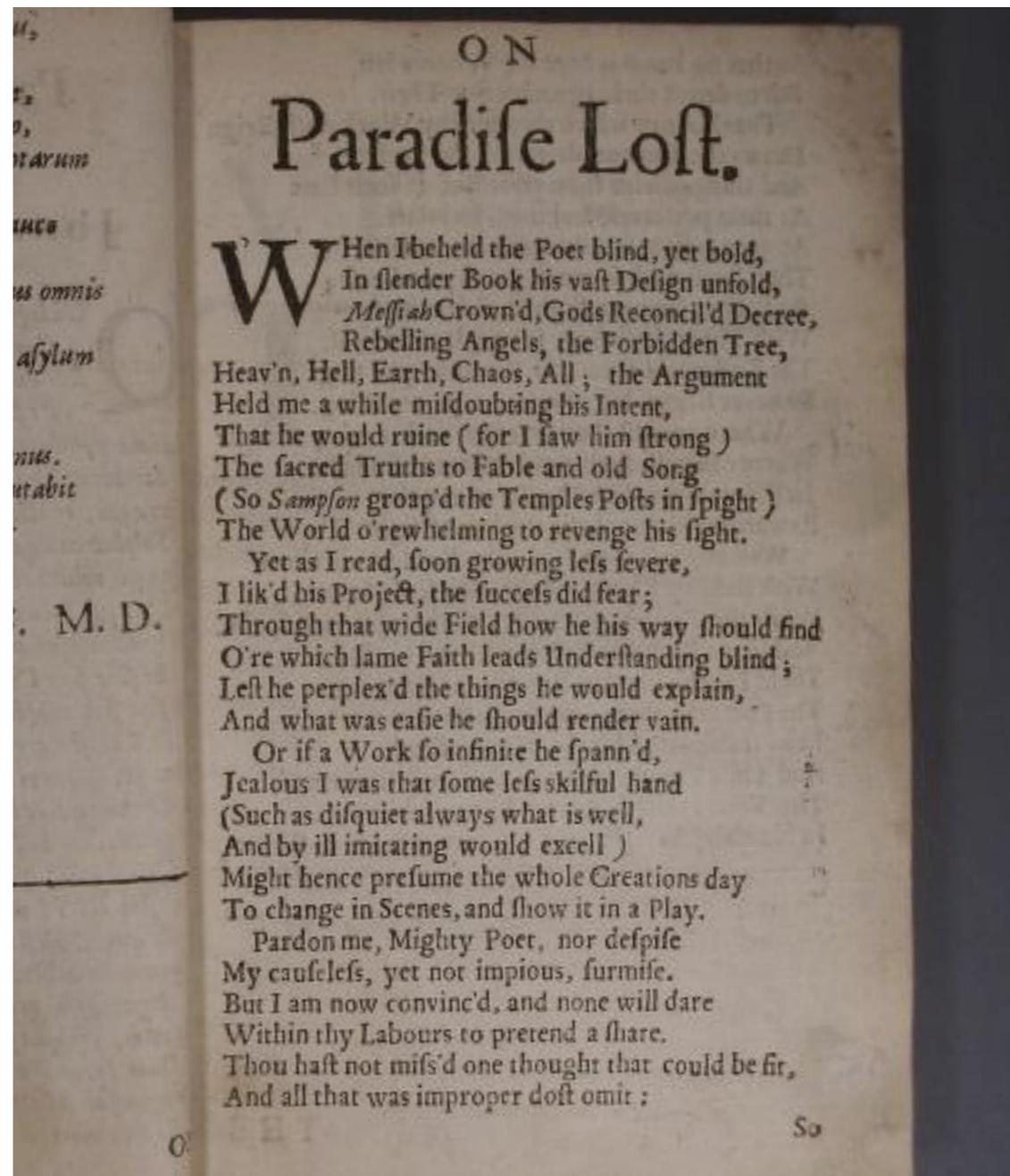
My caufefefs, y €t not impious, furmife.

m;i ?m i n T f onvinc ' d - an <* "one will dare

Within thy Labours to pretend a /hare

InT.nt " Ot mi ^ s ' d one t/,ou § hr fh < 'could be fir

And all that was improper doft omit ; '



www.primeroslibros.com

halperta@gmail.com

www.halperta.com