

# VEP Pipeline Unicode Substitutions

Unlisted characters do not have a direct simple text translation, so they are translated to '@' by the character cleaner. '@' marks can be automatically stripped out or left in place at the discretion of the user. Spacing characters are translated to '|' and automatically stripped out, as they represent visual placement rather than punctuation.

## Greek and Latin Letter Substitutions

Character	Substitution
À	A
Á	A
Â	A
Ã	A
Ä	A
Å	A
Æ	AE
Ç	C
È	E
É	E
Ê	E
Ë	E
Ì	I
Í	I
Î	I
Ï	I
Ð	TH
Ñ	N
Ò	O
Ó	O
Ô	O
Õ	O
Ö	O
×	x
Ø	O
Ù	U
Ú	U
Û	U
Ü	U
Ý	Y
þ	TH
ß	S
à	a
á	a
â	a
ã	a

ä	a
å	a
æ	ae
ç	c
è	e
é	e
ê	e
ë	e
ì	i
í	i
î	i
ï	i
ð	th
ñ	n
ò	o
ó	o
ô	o
õ	o
ö	o
ø	o
ù	u
ú	u
û	u
ü	u
ý	y
þ	th
ÿ	y
Ă	A
ă	a
ć	c
ĥ	h
Ń	N
ő	o
Œ	OE
œ	oe
Š	S
š	s
ţ	t
ů	u
ŵ	w
ÿ	Y
ÿ	Y
ž	z
ı	s
U	V

ḡ	g
υ	v
ζ	z
ϸ	C
Α	A
Γ	G
Δ	D
Ε	E
Θ	TH
Ι	I
Κ	K
Λ	L
Ν	N
Ξ	X
Ο	O
Π	P
Ρ	P
Σ	S
Τ	T
Υ	Y
Φ	PH
Ψ	PS
Ω	O
α	a
β	b
γ	g
δ	d
ε	e
ζ	z
η	e
θ	th
ι	i
κ	k
λ	l
μ	m
ν	n
ξ	x
ο	o
π	p
ρ	r
σ	s
τ	t
υ	u
φ	ph
χ	ch
ψ	ps

ω	o
F	F
A	A
B	B
E	E
M	M
T	T
Y	y
τ	T
ϛ	d
Ϝ	U
“H	H
⋈	A
Ǝ	E
Ɔ	H
Ɔ	H

### Macron Substitutions

Character	Substitution
Ā	A^
ā	a
Ē	E^
ē	e^
Ī	I^
ī	i^
Ō	O^
ō	o^
Ū	u^
ū	u^
Ȫ	y^
ȫ	Y^

### Punctuation and Editorial Mark Substitutions

Character	Substitution
'	'
.	.
i	i
ç	c
£	L
-	@
±	+/-

¶	
.	^
-	-
-	-
—	--
'	'
,	,
“	”
”	”
†	@
‡	@
•	^
...	...
(	(
-	-
--	--
▪	*
◇	@
●	^
{	{
<	(
>	)
÷	/