

# Unblocking #2515

...

Michael Ficarra · 87th Meeting of TC39

# Unicode Tables

68 & 69: property names

Table 68: Non-binary Unicode property aliases and their canonical property names	
Property name and aliases	Canonical property name
General_Category	General_Category
gc	
Script	Script
sc	
Script_Extensions	Script_Extensions
scx	

  

Table 69: Binary Unicode property aliases and their canonical property names	
Property name and aliases	Canonical property name
ASCII	ASCII
ASCII_Hex_Digit	ASCII_Hex_Digit
AHex	
Alphabetic	Alphabetic
Alpha	
Any	Any

70 & 71: property values

Table 70: Value aliases and canonical values for the Unicode property <b>General_Category</b>	
Property value and aliases	Canonical property value
Cased_Letter	Cased_Letter
LC	
Close_Punctuation	Close_Punctuation
Pe	
Connector_Punctuation	Connector_Punctuation

  

Table 71: Value aliases and canonical values for the Unicode properties <b>Script</b> and <b>Script_Extensions</b>	
Property value and aliases	Canonical property value
Adlam	Adlam
Adlm	
Ahom	Ahom
Anatolian_Hieroglyphs	Anatolian_Hieroglyphs
Hluw	
Arabic	Arabic

# Unicode Tables: Why not just normatively refer to Unicode?

68 & 69: property names

- may not want to automatically pull in new property names or aliases

70 & 71: property values

- **do** want to automatically pull in new property values and aliases; no reason to object
- Unicode does not provide a canonical spelling; instead they recommend to do loose matching of names/values (ignores casing, whitespace, hyphens, underscores)
- TC39 previously decided not to do loose matching, so we require a canonical spelling

PR #2515 updates Table 71

# Options for a way forward

1. Approve [#2515](#) via consensus and require all PRs affecting any of these tables to achieve consensus
2. Allow loose matching of Unicode property values and possibly property names
  - remove tables 70/71 and normatively reference Unicode
3. Defer spelling of property values to that used in Unicode spec\*, even though it's explicitly non-canonical
  - remove tables 70/71 and normatively reference Unicode
4. Ask Unicode Consortium to provide canonical spellings for property values and possibly property names
  - remove tables 70/71 and normatively reference Unicode
5. Updates to tables 70/71 use spelling from Unicode spec\* by convention, and do not require consensus

\* NOTE: the names/values are not spelled consistently in the Unicode spec either, but we could define some canonicalisation routine to derive a canonical spelling