

Dictionary line breaking: Thai, Lao, Khmer and Burmese

These scripts put no spaces between words, so no rule can say where a line may end: PeachPDF looks the words up in a word list (ICU's dictionaries) and wraps between them, never inside a syllable. The same Thai (then Lao, then Khmer) paragraph is set in three widths: the words move to the next line whole, whatever the width. The lang attribute does not matter, the script does. The Khmer paragraph also shows real coeng/subjoined-consonant shaping: a subjoined consonant stacks below its base rather than rendering as a separate nominal glyph.

120pt

Age Group	Should Take Action	Should Not Take Action
18-29	88%	12%
30-49	88%	12%
50-69	88%	12%
70+	88%	12%

170pt

A 10x10 grid of blue squares on a light blue background, representing a 100% completion rate.

210pt

A 10x10 grid of blue squares, totaling 100 squares, representing 100% completion.

120pt

ປະເທດລາວຕັ້ງຢູ່ໃນເອເຊຍ
ຕາເວັນອອກສຽງໃຕ້
ນະຄອນຫຼວງວຽງຈັນເປັນ
ເມືອງຫຼວງຂອງປະເທດລາວ
ຂ້ອຍຮັກພາສາລາວ

170pt

ປະເທດລາວຕັ້ງຢູ່ໃນເອເຊຍຕາເວັນອອກ
ສຽງໃຕ້ ນະຄອນຫຼວງວຽງຈັນເປັນ
ເມືອງຫຼວງຂອງປະເທດລາວ ຂ້ອຍຮັກ
ພາສາລາວ

210pt

ປະເທດລາວຕັ້ງຢູ່ໃນເອເຊຍຕາເວັນອອກສຽງໃຕ້
ນະຄອນຫຼວງວຽງຈັນເປັນເມືອງຫຼວງຂອງປະເທດລາວ
ຂ້ອຍຮັກພາສາລາວ

120pt

ខ្ញុំស្រឡាញ់ភាសាខ្មែរណាស់

170pt

ខ្ញុំស្រឡាញ់ភាសាខ្មែរណាស់

210pt

ខ្ញុំស្រឡាញ់ភាសាខ្មែរណាស់

The words as blocks: Burmese

The same idea in the line breaking test font, where every character (a vowel sign or tone mark too) is a square one em wide, and each box holds six of them: a word that does not fit the rest of the line starts the next one, and a syllable is never cut. Burmese is shown this way only (PeachPDF does not yet shape its own stacked/subscript consonants - Khmer's own equivalent is shown above, in real glyphs).

Burmese

