

$$1) ('l', 'o', 'w') : 5$$

$$('l', 'o', 'w', 'e', 's', 't') : 2$$

$$('m', 'e', 'w', 'e', 'n') : 1$$

$$('w', 'j', 'd', 'e', 'n') : 3$$

$$('m', 'e', 'w') : 2$$

Pairs frequencies:-

$$l + o = 7$$

$$o + w = 7$$

$$w + e = 8$$

$$e + s = 2$$

$$s + t = 2$$

$$m + e = 8$$

$$e + w = 8$$

$$e + n = 9$$

$$w + j = 3$$

$$j + d = 3$$

$$d + e = 3$$

$$e + n = 9$$

(l, o, w : 5), (l, o, w, e, s, y : 2),
(m, e, w, er : 6), (w, i, d, er : 3),
(m, e, w : 2)

$$l + o = 7$$

$$o + w = 7$$

$$w + e = 2$$

$$e + s = 2$$

$$s + t = 2$$

$$m + e = 8$$

$$e + w = 8$$

$$w + er = 6$$

$$w + i = 3$$

$$i + d = 3$$

$$d + er = 3$$

$$me = 8$$

$$l + o = 7$$

$$o + w = 5$$

$$w + e = 2$$

$$e + s = 2$$

$$s + t = 2$$

$$me + w = 8$$

$$w + er = 6$$

$$w + i = 3$$

$$i + d = 3$$

$$d + er = 3$$

$$er = 9$$

$$new = 8$$

$$l + o = 7$$

$$o + w = 7$$

$$w + e = 2$$

$$e + s = 2$$

$$s + t = 2$$

$$new + er = 6$$

$$w + i = 3$$

$$i + d = 3$$

$$d + er = 3$$

$$new = 8$$

$$er = 9$$

$$lo = 7$$

$$lo + w = 7$$

$$w + e = 2$$

$$e + s = 2$$

$$s + t = 2$$

$$new + er = 6$$

$$w + i = 3$$

$$i + d = 3$$

$$d + er = 3$$

$$new = 8$$

$$er = 9$$

$$low = 7$$

$$low = 7$$

$$low + e = 2$$

$$e + s = 2$$

$$s + t = 2$$

$$new + er = 6$$

$$w + i = 3$$

$$i + d = 3$$

$$d + er = 3$$

$$new = 2$$

$$er = 9$$

1. ("low": 5), ('low', 'e', 's', 't': 2), ('new', 'er': 6),
('w', 'i', 'd', 'er': 3), ('new': 2)

low + e = 2
e + s = 2
s + t = 2
new + er = 6
w + i = 3
i + d = 3
d + er = 3

Merges

e + s = es
n + e = ne
ne + w = new
l + o = lo
lo + w = low

⇒ Tokenization of "lower": -

('l', 'o', 'w', 'e', 'r')
↑ ↑
'lo', 'w', 'e', 'r'
└───┬───┘ └───┘
 'low' 'er'

= 'low' + 'er'