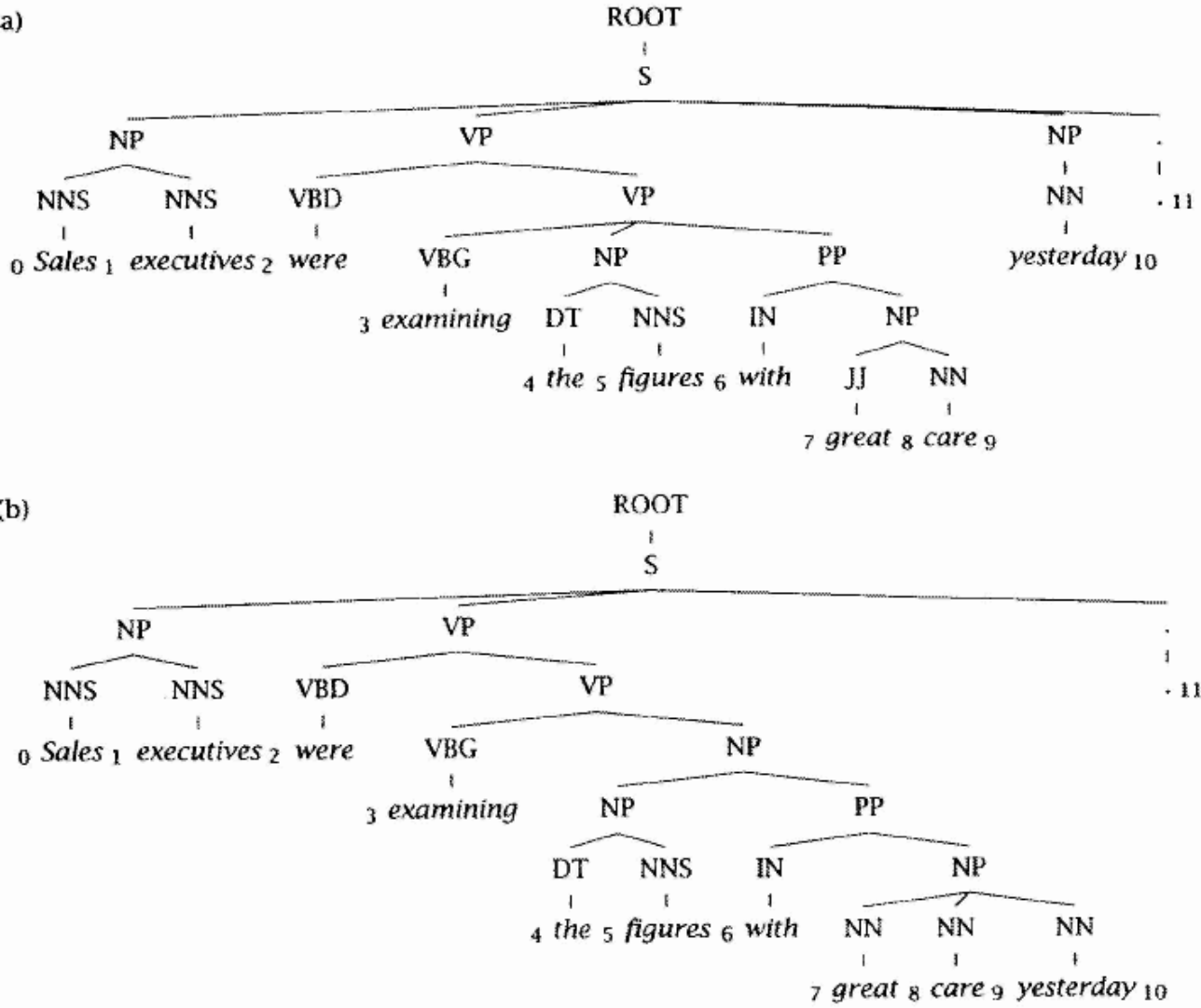

(c) Brackets in gold standard tree (a.):
S-(0:11), **NP-(0:2)**, VP-(2:9), VP-(3:9), **NP-(4:6)**, PP-(6-9), NP-(7,9), *NP-(9:10)

(d) Brackets in candidate parse (b.):
S-(0:11), **NP-(0:2)**, VP-(2:10), VP-(3:10), NP-(4:10), **NP-(4:6)**, PP-(6-10), NP-(7,10)

(e) Precision:	3/8 = 37.5%	Crossing Brackets:	0
Recall:	3/8 = 37.5%	Crossing Accuracy:	100%
Labeled Precision:	3/8 = 37.5%	Tagging Accuracy:	10/11 = 90.9%
Labeled Recall:	3/8 = 37.5%		

Figure 12.6 An example of the PARSEVAL measures. The PARSEVAL measures are easily calculated by extracting the ranges which are spanned by non-terminal nodes, as indicated in (c) and (d) and then calculating the intersection, either including or not including labels while doing so. The matching brackets are shown in bold. The ROOT node is ignored in all calculations, and the preterminal nodes are used only for the tagging accuracy calculation. The starred unary node would be excluded in calculations according to the original standard, but is included here.