I'm not sure exactly where I went wrong here. I used a diff checker for the text itself without any extra space and it found no differences. What really puzzles me is that I got 100% on the test dropbox. I can't really debug because the node and statistics output are the same, maybe I have an extra line return at the end? If someone could explain why it's saying that the test failed, that would be great. This problem has happened on multiple tests.

TEST #3
  bstrees -v ../clean.0 ../display

expected results:
0: "hei"("hei")-
1: "four"("hei")-l "hello"("hei")-r
2: ="five"("four")-l "yksi"("hello")-r
3: "one"("yksi")-l
4: "kaksi"("one")-l "two"("one")-r
5: "kolme"("kaksi")-r ="three"("two")-l ="viisi"("two")-r
6: ="nelja"("kolme")-r
Words/Phrases: 12
Nodes: 12
Minimum depth: 2
Maximum depth: 7

Files yours and expected differ
your results:
0: "hei"("hei")-
1: "four"("hei")-l "hello"("hei")-r
2: ="five"("four")-l "yksi"("hello")-r
3: "one"("yksi")-l
4: "kaksi"("one")-l "two"("one")-r
5: "kolme"("kaksi")-r ="three"("two")-l ="viisi"("two")-r
6: ="nelja"("kolme")-r
Words/Phrases: 12
Nodes: 12
Minimum depth: 2
Maximum depth: 7

Output does not match exactly; the test does not succeed.
Extra newline? Redirect your output to a file (don't cut and paste) and look to see if Maximum depth is the very last line.