Scanner Comparison Report
Manual Scanner vs JFlex Implementation

1. Overview
Both the manual scanner and the JFlex scanner were run against five test files. This report compares their token streams, symbol tables, and error handling.

2. Token Stream Results
Token streams are identical across all five test files. Token types, lexemes, line numbers, and column numbers match exactly.

Test File	Manual Tokens	JFlex Tokens	Errors	Match
test1.lang	268	268	0	✓ Yes
test2.lang	191	191	0	✓ Yes
test3.lang	112	112	0	✓ Yes
test4.lang	56	56	10	✓ Yes
test5.lang	53	53	0	✓ Yes

3. Error Detection — test4.lang
Both scanners detected the same 10 lexical errors.

Line	Lexeme	Manual	JFlex + ErrorHandler
6	invalidVar	ERROR token	Invalid Identifier
9	@	ERROR token	Invalid Character
12	3.1234567	ERROR token	Malformed Literal
15	"this string never ends;	ERROR token	Unclosed String
18	"bad \x escape"	ERROR token	Invalid Escape
21	Abcdefghijklmnopqrstuvwxyz_extra	ERROR token	Identifier Too Long
24	'<newline>	ERROR token	Unclosed Char
27	"	ERROR token	Malformed Literal
30	'ab'	ERROR token	Malformed Literal
33	~	ERROR token	Invalid Character

4. Symbol Table Differences

The only consistent difference between the two scanners is how they populate the symbol table. This is a driver-level design choice, not a scanner error.

Test File	Manual Rows	JFlex Rows	Reason
test1.lang	57	11	Manual adds every occurrence; JFlex adds unique only
test2.lang	45	7	Same reason as above
test3.lang	27	14	Same reason as above
test4.lang	8	8	Identical — only valid identifiers added
test5.lang	10	2	Same reason as above

5. Key Differences Summary

Aspect	Result	Detail
Token stream	IDENTICAL	Same tokens, same order, same positions
Token count	IDENTICAL	268/191/112/56/53 — all match
Error detection	IDENTICAL	Same 10 errors caught in test4.lang
Error reporting	DIFFERENT	JFlex gives categorised messages; manual gives bare ERROR tokens
Symbol table content	DIFFERENT	Manual: all occurrences. JFlex: unique identifiers only
Comments stripped	IDENTICAL	Both correctly remove ## and #* ... *# comments

6. Conclusion

The JFlex scanner is functionally equivalent to the manual scanner. Token streams, token counts, and error detection are identical across all five test files.

The two differences — symbol table population and error reporting format — are both driver-level choices, not scanner defects. The JFlex implementation is improved in error reporting thanks to the ErrorHandler integration.