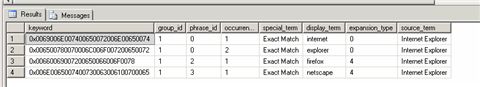
DM\_FTS\_PARSER takes a full text query and breaks it up using the word breaker rules, applies stop lists (more on them later), and any configured thesaurus. This is essential in the first step of diagnosing when users are complaining because there queries aren’t working. Often this is due to, a word not breaking as expected, use of noise words that exist in the stop list or thesaurus replacing  or substituting words.

You call the function using the same query string as you would use normally with a CONTAINS statement, along with a language, a stop list and where the search should be accent sensitive.

SELECT \*   
FROM sys.dm\_fts\_parser ('FORMSOF( THESAURUS, "Internet Explorer")', 2057, 0, 0)

This  returns the following,



You can see that in my thesaurus I have added substitution elements for Internet Explorer or firefox and netscape.

The following query ,

SELECT \*   
FROM sys.dm\_fts\_parser ('multi-million', 2057, 0, 0)

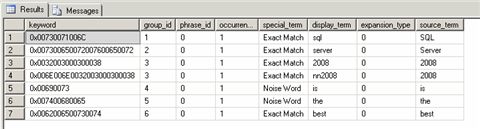
Returns the following showing how the word breaking as broken the word up but also maintained the combined word.

http://sqlblogcasts.com/Photos/blogimages/images/7559/480x480.aspx

Finally

SELECT \*   
FROM sys.dm\_fts\_parser ('SQL OR Server OR 2008 OR is OR the OR best', 2057, 0, 0)

Returns the following which nicely indicates which words are noise words but also that numbers are searched as numbers and text. Note the nn prefix.



And finally finally, the query about c++, c# etc.

SELECT \*   
FROM sys.dm\_fts\_parser ('C or c or C++ or c++ or C# or c#', 2057, 0, 0)

Returns the following, which shows what you need to put in to get an exact search on c++, or c#. Capitalise the C. What’s also interesting is that C, C++ both relate to C as well but C# doesn’t, which means it C is removed from the noise word then C++ would return any document containing the word C.

