

REMARK

To The Paper Of

Joseph E. Sullivan

by

Robert A. J. Gildea

We have in the DOTSYS program some things that you have mentioned. The special character which logically terminates part of a word is present as one of the features that can be inserted into the input stream and force the separation at that part, so that a contraction will not cross that boundary. There are two such things; one is sort of a literary parenthesis which says, "under any circumstances please contract this group of letters", and the other is, "this is a separator; do not go beyond this in the search for a group of characters". So both of those are present as controls for the Braille editor to put into the input stream. The other thing you mentioned about the 50,000 words that Mr. Vinding brought up . . . Mr. Sullivan mentioned a self-check feature. That self-checking feature has 5,808 problem words, the more difficult ones of English, American version. And that is an input control, you can just put in a control card that says "run self-check". It will go through the 5,808 - nearly 6,000 words comparing and counting and printing out the errors for you, numbering them, and at the end you look at the number. If it is the familiar 138, or whatever it is, that you know, then you figure the modification you made in the program has not affected its translating capability.

We have not taken out all the errors that are flagged by that self-checking approach, simply because it is more work than we want to do; it is unnecessary to take them out. We may lose something in some cases if we take them out.

Some of the things we have in there are trivial, such as Joe Sullivan mentioned, and that was do and do, things of that nature, where it is a semantic difference, and there is no way for any economic computer program to delve into the meaning. We know we could do that with some transformational grammar approach, or something, to dive into the meaning of the sentence, come up with not only the meaning, the syntax, but then go after the semantics and get that but that's not for production Braille translation.