Brain and Language **68,** 218–224 (1999)
Article ID brln.1999.2085, available online at http://www.idealibrary.com on IDEAL®

Decomposition: To What Extent? The Case of Turkish

Ayşe Gürel

McGill University, Montréal, Québec, Canada

It has been proposed that in agglutinative languages, lexical access of morphologically complex words must involve decomposition rather than full listing (Frauenfelder & Schreuder, 1992; Hankamer, 1989). We tested this proposal in Turkish using a simple lexical decision task. Results show that multimorphemic words that consist of frequent affixes are processed as fast as monomorphemic words. This finding suggests that in languages with rich morphology, not all multimorphemic words are accessed in a decomposed form. To the extent that morphemes are in frequent use, they may induce whole-word rather than decompositional lexical access. © 1999 Academic Press

Key Words: Turkish; lexical access; morphological decomposition; whole-word access.

INTRODUCTION

Lexical access and representation of complex words has widely been discussed in the psycholinguistic literature. Existing models of lexical access of multimorphemic words range from morphological decomposition (Taft & Forster, 1975) to full-listing (Butterworth, 1983). While the decompositional account assumes that a morphologically complex form is parsed into its constituent morphemes prior to lexical access, the full-listing view maintains that the morphological structure of a complex form has no independent representation, suggesting that no parsing is involved in word recognition. There are also hybrid models which include features of both decomposition and full-listing models (Caramazza, Laudanna, & Romani, 1988; Frauenfelder &

This research was in part supported by an MCRI grant from the Social Sciences and Humanities Research Council of Canada to Gonia Jarema, Eva Kehayia, and Gary Libben. I thank all the members of the Mental Lexicon Project, especially Eva Kehayia, Chris Grindrod, Gerald Rosenau, and Kyrana Tsapkini for their invaluable help in the design and the statistical analysis of this experiment. I also thank the conference participants in Edmonton for their helpful suggestions.

Address correspondence and reprint requests to Ayşe Gürel, Department of Linguistics, McGill University, 1001 Sherbrooke Street West, Montréal, Québec, H3A 1G5 Canada. E-mail: agurel@po-box.mcgill.ca.

0093-934X/99 $30.00

Schreuder, 1992). Among the hybrid models, Augmented Addressed Morphology (AAM) (Caramazza et al., 1988) postulates that an input can activate simultaneously both a whole-word representation and constituent morphemes. When a word is familiar to the subject, whole-word representation is activated. When the word is novel, however, morpheme activation takes place. The Morphological Race Model (MRM) proposed by Frauenfelder and Schreuder (1992) also assumes the existence of two routes which are in competition. However, in their view, even familiar words can be accessed through decomposition depending on factors such as transparency and frequency. According to this model, while the parsing route wins in the recognition of transparent low-frequency words, the direct route wins in the recognition of high-frequency opaque words.

The above-mentioned models have been tested using only a few languages (for the most part English). However, it has recently been acknowledged that cross-language studies, particularly those using languages with a rich morphology (such as Turkish and Finnish), allow for experimental confirmation of the proposed models of lexical access (Laine, Niemi, Koivuselkä-Sallinen, Ahlsén, & Hyönä, 1995). It has been suggested that due to the storage efficiency in highly inflected, agglutinative languages like Turkish, lexical access of morphologically complex words must involve decomposition rather than full-listing (Hankamer, 1989). A similar claim has come from Frauenfelder and Schreuder (1992) who suggest that in Turkish, the morphological parsing route rather than the direct access route must win the race in the analysis of most complex words. However, they also predict that depending on the frequency of the root-plus-affix combination, a word can be recognized by the direct route.

In order to test the above predictions, we investigated word recognition of morphologically simple and complex words in Turkish using a simple lexical decision task measuring reaction time (RT). In this study, we also looked at whether stems that include pseudomorphemes (i.e., units which are homographic and homophonous with real morphemes) are more difficult to process than simple nondecomposable stems. There are many Turkish words (i.e., stems) that can induce ambiguous parsings. That is, assuming a left-to-right parsing for Turkish (Hankamer, 1989), an input stem that is presented in isolation can have several alternative parsings. For example, the word *dilim* is ambiguous in the absence of any contextual information. Thus, depending on the access strategy, it can have different readings: it means ''slice,'' if accessed as a whole or it means ''my tongue,'' if parsed as *dil* + *im* (''tongue'' + first-person singular possessive). We assume that if every morphemelike representation in a word is activated in word recognition, then the processing will take longer and this, in turn, will lead to longer RTs for these pseudomorphemic items.

Given these features of the language, we investigated the following questions: (1) to what extent does lexical access of multimorphemic words in

Turkish involve morphological decomposition? and (2) will all possible substrings of a word be parsed in word recognition? We hypothesize that (1) if decomposition takes place during word recognition, we anticipate longer RTs for all multimorphemic words and (2) if every possible morpheme or morphemelike structure is activated in word recognition, we anticipate longer RTs not only for multimorphemic but also for pseudomorphemic items.

METHOD

Participants

Thirty-one native speakers of Turkish, with an average age of 26 years (range 18–36) and an average number of 19 years of education (range 12–25) participated in this study.

Procedure

Subjects were given a simple visual lexical decision task run on a Power Macintosh using PsyScope 1.1. Subjects saw a string of letters presented on the computer screen and were asked to press the ''yes'' key if they recognized the item as a word of Turkish and the ''no'' key if they did not. The main experiment was run in a single test of 576 trials. A practice trial of 10 stimuli preceded the main test. Each stimulus was preceded by a mask (######) lasting 150 ms followed by a 200-ms interval until the target items appeared on the screen. The item remained on the screen until the subject pressed ''yes'' or ''no.''

Stimuli

The stimuli comprised 130 words, 273 nonwords, and 173 fillers. The nonwords were constructed by changing the first phoneme of the real-word stimuli. Nonwords included both simple and inflected types, affixed with a legal suffix of Turkish. Fillers consisted of verbs, adjectives, and adverbs, both simple and inflected.

The experimental stimuli comprised monomorphemic and multimorphemic items. Monomorphemic items included nondecomposable and pseudomorphemic items.

Nondecomposable items (NDC) are words that cannot be decomposed in any way (e.g., *pencere*, ''window'').

Pseudomorphemic items were of three types. (1) pseudostem (PS): this group included words such as *dalga* (''wave''), which consist of a meaningful stem plus a syllable which has no meaning in Turkish. For example, *dal* (''branch''), the first syllable of the word, is followed by a meaningless syllable, *ga*. (2) Pseudostem-stem (PSS): this category included pseudomorphemic items which appear to contain two stems. For example, the first and second syllables of the word *bakkal* (''grocery'') are also meaningful stems in Turkish: *bak* (''look'') and *kal* (''stay''). (3) Pseudostem-affix (PSA): this category consisted of pseudostems which have a stem followed by a legal suffix in Turkish. These were homographic and homophonous with the possessive suffix in Turkish [e.g., *dilim* (''slice''), *dil* (''tongue''), and *im* (the first-person singular possessive)].

Multimorphemic items contained one- and two-suffix words inflected for case (ablative, locative) and number [e.g., stem-ablative (S-AB): *deprem-den* (''from the earthquake''); stem-locative (S-LOC): *masa-da* (''on the table''); stem-plural (S-PL): *emir-ler* (''orders''); stem-plural-ablative (S-PL-AB): *oda-lar-dan* (''from the rooms''); and stem-plural-locative (S-PL-LOC): *resim-ler-de* (''in the pictures'')]. These morphemes differ in frequency in both

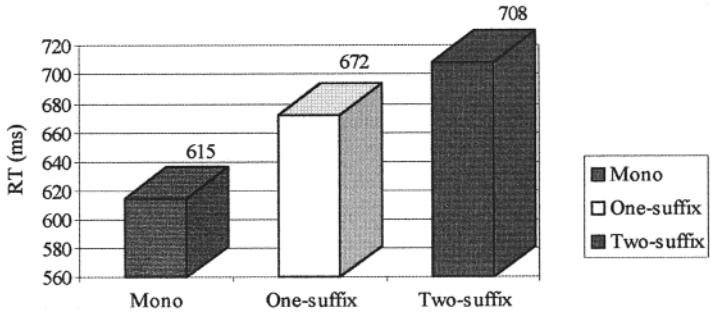

FIG. 1. Mean RTs for one-suffix, two-suffix, and monomorphemic words.

written and spoken Turkish: the ablative suffix has the lowest frequency. This is followed by the locative suffix. The plural morpheme has the highest frequency (Pierce, 1960).

Multimorphemic and monomorphemic items were matched only for stem frequency. In the frequency count of Turkish (Pierce, 1960), the stem and the morpheme frequency are given separately. Therefore, the surface frequency was not available. Nondecomposable items were matched with one-suffix words for length. Length here can be defined in terms of either the number of syllables or the number of letters. In either case, one-suffix words were matched with nondecomposable items in length. Nondecomposable words of three to four syllables in Turkish are generally of low frequency. Therefore matching these items with one-suffix words for frequency was not possible. Pseudomorphemic words were inevitably shorter than nondecomposable monomorphemic items.

Results and Discussion

The error rate across categories ranged from 0.5 to 3.4%. The RTs that were ± 2 standard deviations above and below each group mean were deemed outliers and thus were eliminated from further analysis. Errors and outliers constituted 6.8% of total responses. The results show that as a whole, multimorphemic words (687 ms) yield longer RTs than monomorphemic words (615 ms) ($p < .0001$). A repeated-measures ANOVA revealed that overall there is a significant difference among one-suffix, two-suffix, and monomorphemic words [$F(2, 30) = 24.11$, $p < .0001$]. Post hoc comparisons using the Fisher's LSD method showed that monomorphemic words yield significantly shorter RTs than one-suffix ($p < .05$) and two suffix words ($p < .005$) (see Fig. 1).

Comparisons of the monomorphemic group to each category of the one-suffix group reveal an important effect of the suffix frequency. A repeated-measures ANOVA for these comparisons revealed a significant difference [$F(3, 30) = 9.8$, $p < .0001$]. As shown in Fig. 2, a Fisher's LSD post hoc analysis showed that the words with the ablative suffix (S-AB) yield significantly longer RTs than monomorphemic words ($p < .005$). The interesting finding is that the difference between words with the locative case marker (S-LOC), the second most frequent suffix, and monomorphemic words is not statistically significant. Furthermore, there is no significant difference in

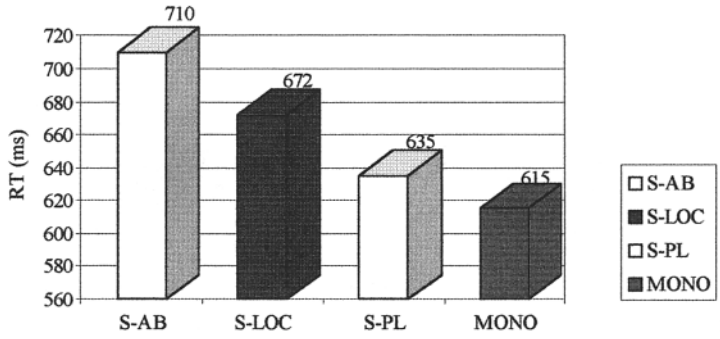

FIG. 2. Mean RTs for each category in the one-suffix group and monomorphemic words; S-AB vs MONO, $p < .005$; S-LOC vs MONO, ns; S-PL vs MONO, ns.

RT for words with the frequent plural suffix (S-PL) and monomorphemic words. That is, words with the plural morpheme are accessed as fast as monomorphemic words.

The individual analysis within one-suffix words shows an effect of frequency of suffixes in lexical access. A repeated-measures ANOVA revealed a significant difference among the categories in this group [$F(2, 30) = 7.01$, $p < .005$]. As Fig. 3 shows, words with the ablative suffix, which has the lowest frequency, yield the longest RT. This is followed by the words with the locative suffix. The words which contained the plural morpheme, however, yield significantly faster RTs. A Fisher's LSD post hoc analysis showed that words with the ablative suffix yield significantly longer RTs than the words inflected with the plural morpheme ($p < .05$).

Within monomorphemic words, a repeated-measures ANOVA yielded a significant difference [$F(3, 30) = 7.9$, $p < .0001$]. However, a Fisher's LSD post hoc analysis did not show any significant difference between the nonde-

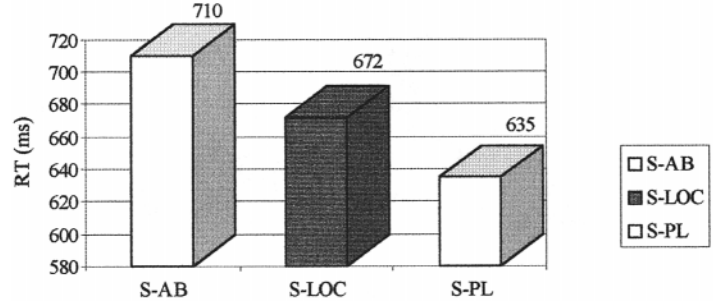

FIG. 3. Mean RTs for each category in the one-suffix group; S-AB vs S-LOC, ns; S-AB vs S-PL, $p < .05$; S-LOC vs S-PL, ns.

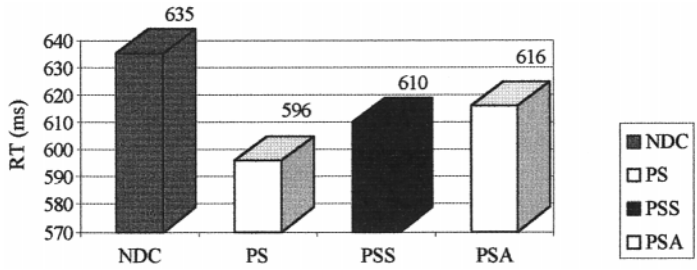

FIG. 4. Mean RTs for nondecomposable (NDC) and pseudomorphemic words (PS, PSS, and PSA).

composable group and any of the categories in the pseudomorphemic group (Fig. 4). In this analysis, the only difference that was close to reaching significance was between NDC and PS ($p = .07$). However, it was not statistically significant.

As mentioned earlier, in a left-to-right parsing in Turkish, there may be several alternative parsings in the lexical access of certain words and this is expected to cause a delay in processing, and, in turn, longer RTs. However, it seems that not every ''possible'' substring of a word is parsed. If this had been the case, we would have obtained longer RTs for pseudomorphemic items.

A comparison among multimorphemic words reveals a significant interaction between the number of suffixes and the speed of word recognition. We find that overall, two-suffix words yield significantly longer RTs than one-suffix words ($p < .05$). A repeated-measures ANOVA for the multimorphemic group yielded a significant difference [$F(4, 30) = 6.03$, $p < .0005$]. However, as can be seen in Fig. 5, the two-suffix group (both S-PL-AB and S-PL-LOC) are accessed as fast as one-suffix words (S-AB and S-LOC). A Fisher's LSD post hoc analysis revealed that the only significant difference is between one-suffix words with the plural morpheme (S-PL) and two-suffix words with the plural plus ablative morpheme (S-PL-AB) ($p < .05$). This suggests that the plural morpheme in two-suffix words does not contribute any extra processing load to the access of these words; it is accessed as part of the stem.

CONCLUSION

This study has shown that not all multimorphemic words are accessed in a decomposed form in Turkish. Words with frequent suffixes seem to be accessed through a whole-word access procedure. Depending on the frequency of the suffix, a word can be accessed via the direct access route or the parsing route. The higher frequency of a suffix makes processing easier and faster. It appears, then, that in a highly inflected language like Turkish,

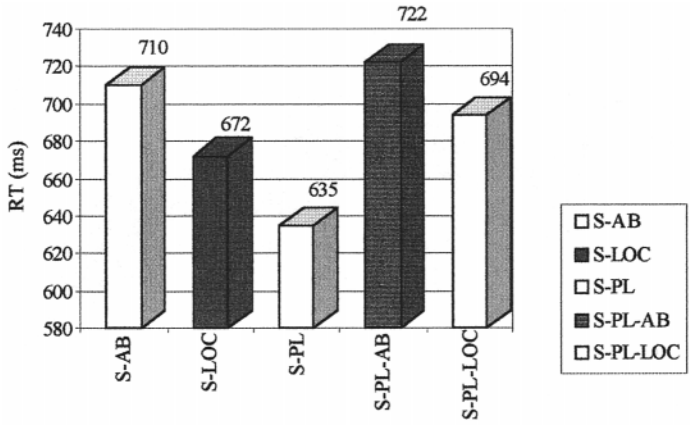

FIG. 5. Mean RTs for multimorphemic words; S-AB vs. S-PL-AB, ns; S-AB vs S-PL-LOC, ns; S-LOC vs S-PL-AB, ns; S-LOC vs S-PL-LOC, ns; S-PL vs S-PL-AB, $p < .05$; S-PL vs S-PL-LOC, ns.

the recognition of morphologically complex words is not as ''costly'' as it may be in languages with little inflection (such as English), at least for words that contain frequent suffixes. In other words, lexical access in languages with complex morphology involves the direct route where possible in order to save time in processing. This suggests that the parser in Turkish needs to be much more effective than the parser in English in order to handle the complexity of word forms during lexical access.

REFERENCES

Butterworth, B. 1983. Lexical representation. In B. Butterworth (Ed.), *Language production* (Vol. 2). New York: Academic Press.

Caramazza, A., Laudanna, A., & Romani, C. 1988. Lexical access and inflectional morphology. *Cognition,* **28,** 297–332.

Frauenfelder, U. H., & Schreuder, R. 1992. Constraining psycholinguistic models of morphological processing and representation: The role of productivity. In G. E. Booij & J. van Marle (Eds.), *Yearbook of morphology 1991.* Dordrecht: Kluwer.

Hankamer, J. 1989. Morphological parsing and the lexicon. In W. Marslen-Wilson (Ed.), *Lexical representation and process.* Cambridge, MA: The MIT Press.

Laine, M., Niemi, J., Koivuselkä-Sallinen, P., & Hyönä, J. 1995. Morphological processing of polymorphemic nouns in a highly inflected language. *Cognitive Neuropsychology,* **12**(5), 457–502.

Pierce, J. 1960. *A frequency count of Turkish words.* Ankara: Milli Egitim Mudurlugu.

Taft, M., & Forster, K. I. 1975. Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning and Verbal Behavior,* **14,** 638–647.