A WAY OF TESTING THE PRODUCTIVITY OF WORD FORMATION RULES (WFRs)?

ELŻBIETA GÓRSKA

Adam Mickiewicz University, Poznań

The purpose of this paper is to provide a support for the idea that the productivity of a WRF should be viewed not only as a diachronic but also as a synchronic process (cf. Aronoff 1978). The long lasting debate on whether word formation (WF) should be treated as a dynamic phenomenon need not be repeated here (cf., e.g. Hockett 1954; Pennanen 1972; Beard 1974). Having assumed that WF is not only a field of study but also a dynamic process, this paper will concentrate on the possibilities of measuring the productivity in WF[1]. In particular, a variant of the lexical decision task technique (cf. Aronoff 1980), will be put forward as a means for testing the productivity of a WFR.

The methodology of testing the productivity of WF processes is certainly not a very well developed field. This seems to be at least particularly due to the debate mentioned above which has often resulted in WF being treated as an a posteriori phenomenon — a closed system (cf., e.g., Heinz 1961). Only recently, because of the view that WF is also a synchronic process (cf. Aronoff 1980) and that there exists some kind of native speaker's WF competence (cf. Pennanen 1972) has the problem of testing the productivity of WFRs arisen.

In Polish linguistics, mainly under the influence of Dokulil's (1979) theory of onomasiological categories, experiments testing "motivation" of derived forms have been carried out (e.g. Nagórko-Kufel 1977). In them, the main emphasis was placed on the degree of dependence of derived forms on their formal and semantic relations to the bases. The informants were asked to

---

[1] Productivity is understood here as in Aronoff (1976 : 37): "likehood of being a word in the speaker's active vocabulary".

name a word or words related formally and semantically to a given derived form. The results of these experiments shed some light on the factors that may influence WF processes. While analyzing the answers given, it appeared that the formal relations were by far more important to the informants than the semantic ones. The experiments do not, however, reveal anything about the degree of productivity of WFRs.

A test to measure the productivity of WFRs in English has been presented by Aronoff (1980). There is no place here for describing the test in detail. For our purpose, however, it suffices to say that it is rooted in the lexical decision task technique developed by cognitive psychology and that the informants are given three different types of items: a. *words* — actual words in the language; b. *possible words* — these items do not occur in the language but the bases from which they are derived do; c. *non-words* — neither these items nor the bases occur in the language.

The test carried out by the present author is similar to the test presented by Aronoff (1980). It differs in that all the words that are presented to the informants do exist in the language analyzed here. Based on the well known fact that every native speaker of a language has a different dictionary (the latter is understood as the lexicon of an individual speaker), it was assumed that the three questions given by Aronoff can be asked with reference to *existing* words of a language. It was expected that, comparing two WFRs that operate at least partly on the same type of bases, the one which is more productive will have a higher percentage of "possible but non-occurring" words than the one which is less productive. The test consisted of two lists of 38 Polish deverbal nouns each (all cited in the *Index a tergo*): the first included nouns such as *składacz* "type-setter" and the second, nouns such as *składak* "folding canoe" (the verbal bases in each list were the same, here: *składać* "to set together, fold"). There were two groups of informants: 28 secondary school pupils of LO at Czarnkowo and 22 secondary school pupils of LO XXIV in Warsaw[2].

A detailed language internal analysis of the data is presented in the author's MA thesis (1978). Simplifying matters for the purpose of this paper, here are some of the most relevant points. It has been determined that within the framework of Aronoff (1976), the WFR that forms deverbal nouns such as *składacz* operates on a uniquely specified class of bases, i.e. those imperfect verbs with the $+a+$ ending in the infinitival form. The meaning of derived noun, although it is a compositional function of the meaning of the base, i.e. "sb that X_V", "sth that X_V", cannot be predicted. On the other hand, the

WFR that creates the nouns *składak* etc., is totally different: neither the category of the base (the nouns in $+ak$ are formed from nouns, verbs and adjectives), nor the meaning of derived words can be predicted. On the basis of the analysis, it was assumed that the WFR for *składacz* is more productive than the one for *składak*. Here, the test for productivity of WRFs was applied[3].

As has already been mentioned, the author assumes that the differences in question should be reflected in the lexicon of Polish. The informants were given two lists of Polish nouns described above and were asked to identify every word according to the following criteria:[4] A word is:

a. actual FOR YOU, i.e. you use that word and you know what it means, or somebody had used that word and you knew what it meant.

b. possible though non-occurring, ACCORDING TO YOU, i.e. you have never yet used or heard that word but you could use that particular word to name something or somebody, or somebody else could use it.

c. impossible.

The table below illustrates the results of the test:

	SKŁADACZ			SKŁADAK		
total	actual	possible	impossible	actual	possible	impossible
1900	645	636	619	649	472	779
100%	34%	33%	33%	34%	25%	41%

One can easily notice that the percentage of "actual" words is identical for both WFRs (34%), while those of "possible" and "impossible" items differ considerably. It seems that these differences reveal the regularities that exist within the dictionaries of native speakers — the degree of productivity of the rules in question is not the same. The fact that there is a higher percentage of "possible" items of the *składacz* type (33%), than of the *składak* type (25%) seems to confirm the assumption that the WFR for *składacz* etc., is more productive than the one that creates nouns such as *składak*.

A few tentative conclusions can be drawn at this point:

— different treatment of existing words in a language reflects the produc-

---

[2] These tests were carried out in 1980. Two years ago I presented a preliminary version of the test to my collegues — the students of the English Department, Warsaw University. Although the basic assumption of the analysis (see below) was then confirmed, the number of informants (15) was too small to make a comparison.

[3] See fn. 2.
[4] There were two ways of checking whether the test was completed randomly:
— in each list a noun that is used very often in every day speech was chosen as a "control word" *otwieracz* "opener" for the first and *zmywak* "wash clout" — for the second list. If an informant marked the noun as "possible" or "impossible" the test was rejected.
— also, two „possible though non-occurring" items were added to both lists (in the first: *suwacz from *suwać* "push" and *zderzacz — *zderzać* "collide", and in the second *przedłużak — *przedłużać* "lengthen" and *otwierak — *otwierać* "open"), and if these were marked as "actual", the test was considered invalid.

tivity of WFRs. The fact that native speakers qualify some of the actual words as "possible" in a systematic way, proves that every native speaker of a language can form and accept new words by applying WFRs of his language. Hence, there exists some kind of WF competence.

— the results of the test contradict a possible flaw of the main assumption accepted above that might have already occurred to the reader, namely, that is not memory that is tested (or, at least, not only memory). If only memory were tested then, most probably, the words would be classified as either "actual" or "impossible" with none falling into the category "possible", or, if some were labelled as being "possible", then the choice of these words would not exhibit the regularities discussed above.

Finally, it seems useful to distinguish between two types of potential bases (or, whatever we call them) in WF. They are:

1. potential for a given native speaker. Since the speaker considers some existing words in the language as "possible", he establishes his own potential bases for a given WFR, i.e. if the base fulfills the conditions of that rule (see the regularities revealed by the test).

2. potential for a language as a whole. This can be assumed provided that (1) is true, i.e. that the conditions for the application of a given WFR which are at work in (1) also hold for (2). Thus, if we can find two WFRs that operate at least partly on the same class of bases, and if one of the rules has more transparent conditions on its application (e.g. unique specification of the base; meaning of a derived form — a semantic function of the base, as in *składacz*) than the other one (cf. *składak*), then the test presented here should reveal that the amount of potential bases of the first type is higher for the former than for the latter WFR. From this the degree of productivity of the two WFRs in the language as a whole can be established. It can be expected that the WFR with clear conditions would also have a higher number of potential bases of the second type than the one for which no conditions can be found — more items would enter the language by the application of the former process than by the latter (in our case, we except a higher number of the nouns such as *składacz* than *składak* to appear in Polish)[5].

There is no doubt that in order to accept the assumptions mentioned above, the test presented here needs a far more detailed study on a larger scale. This seems also to be the only way to get rid of the question mark placed in the title of this paper.[6]

[5] The existence of the bases of the 2nd type can be further confirmed by using the test for productivity developed by Aronoff (1980).

[6] Let me add that the same test was applied to a totally different piece of data from Polish: 47 nouns such as *trener* "trainer" (i.e. borrowings — mainly from English) which

LIST I

1. *nagarniacz* "mechanism of a harvester"
2. *zgarniacz* "scraper"
3. *krajacz* "cutter"
4. *nabijacz* "loader"
5. *odbijacz* "printer"
6. *podbijacz* "(archaic) conqueror"
7. *ubijacz* "beetle, rammer"
8. *wybijacz* "a worker of a foundry practice"
9. *nawijacz* "wider"
10. *zawijacz* "a person that packs"
11. *odchylacz* "deflector"
12. *przecinacz* "cutter"
13. *ścinacz* "radiology: chipper"
14. *wycinacz* "a worker of weaving industry"
15. *odcinacz* "forestry: a worker that cuts off branches"
16. *zabieracz* "a worker of mining industry"
17. *odbieracz* "receiver"
18. *wybieracz* "a person that selects sth"
19. *potrząsacz* "a tool used for mechanical handling"
20. *wstrząsacz* "shaker"
21. *czesacz* "weaving: a worker that weaves wool"
22. *rozwiertacz* "driller worker"
23. *chwytacz* "machine parts: catcher"
24. **suwacz*
25. *nakluwacz* "needle point of compasses"
26. *pływacz* "bladderwort"
27. *podbieracz* "pick up (in a combine harvester)"
28. *pomywacz* "dish-washer (worker)"
29. *zmywacz* "paint remover"
30. *mazacz* "a person that scribbles"
31. *wytaczacz* "a worker of machine industry"
32. *oczyszczacz* "cleaner"
33. *rzezacz* "(archaic) sculptor"
34. *mieszacz* "mixer (worker)"

LIST II

1. *nagarniak* "a tool used in mining"
2. *zgarniak* "scraper bucket"
3. *krajak* "a kind of knife used in carpentry"
4. *nabijak* "tamping stick"
5. *odbijak* "scaling hammer"
6. *podbijak* "railways: tamper"
7. *ubijak* "rammer"
8. *wybijak* "a tool: drift"
9. *nawijak* "reeling apparatus"
10. *zawijak* "curling die"
11. *odchylak* "back iron"
12. *przecinak* "chipper"
13. *ścinak* "chisel"
14. *wycinak* "blanking die"
15. *odcinak* "cutting off die"
16. *zabierak* "machine tools: driver"
17. *odbierak* "current collector"
18. *wybierak* "selector"
19. *potrząsak* "mining: shaker"
20. *wstrząsak* "a part of machine for making matches"
21. *czesak* "a part of weaving machine"
22. *rozwiertak* "reamer"
23. *chwytak* "grab (bucket)"
24. *suwak* "slider, zip"
25. *nakluwak* "foundry practice: vent wire"
26. *pływak* "swimmer"
27. *podbierak* "a kind of fishing net"
28. *pomywak* "wash clout"
29. *zmywak* "a dish clout"
30. *mazak* "a thick painting brush"
31. *wytaczak* "boring tool"
32. *oczyszczak* "a part of assembly belt"
33. *rzezak* "machine parts: knife"
34. *mieszak* "machine parts: mixor"

were assumed to form a class of bases for Nomina Agentis and Nomina Instrumenti (marked in Plural as *trenerzy, trenery* respectively). The assumption that both WFRs are equally productive (there are no formal conditions for the application of the processes in question), has been confirmed by the results of the test, i.e. the number of "possible though non-occurring" items is almost the same (Nomina Agentis — 14%, Nomina Instrumenti — 15%). This, however, was only a preliminary analysis (27 secondary school children) and no definite conclusion can be drawn here.

35. *zderzacz
36. *wydłużacz* "metallurgy: elongator"
37. *przedłużacz* "extension rod"
38. *pełzacz* "creeper"
39. *składacz* "type-setter"
40. *otwieracz* "opener"

35. *zderzak* „bumper"
36. *wydłużak* "tools: fuller"
37. *przedłużak
38. *pełzak* "creeper, amoeba"
39. *składak* "folding canoe"
40. *otwierak

REFERENCES

Aronoff, M. 1976. *Word formation in generative grammar*. Cambridge Mass., : The MIT Press.

Aronoff, M. 1980. "The relevance of productivity in a synchronic description of WF". In Fisiak, J. (ed.). 1980. 71—82.

Beard, R. E. 1974. "Word formation: the anarchy of the arts". Revision of the paper presented at the 19th Annual Meeting of the International Linguistic Association, New York 9—10 March, 1974. (A microcard from Educational Resources Information Center Reports. Arlington, Virginia.)

Dokulil, M. 1979. *Teoria derywacji*. Wrocław: Ossolineum.

Fisiak, J. (ed.). 1980. *Historical morphology*. The Hague: Mouton.

Górska, E. 1978. *Nominalization in Polish and English. Sample analyses*. Unpublished MA thesis, University of Warsaw.

Heinz, A. 1961. "Fleksja a derywacja". *Język Polski XLI*. 343—54.

Hockett, C. F. 1954. "Two models of grammatical description". *Word 10*. 210—31.

Index a tergo do Słownika j. polskiego pod redakcją Witolda Doroszewskiego (Grzegorczykowa, R. and J. Puzynina (eds). 1973. Warszawa: Państwowe Wydawnictwo Naukowe.)

Nagórko-Kufel, A. 1977. "O eksperymencie w badaniach słowotwórczych". *BPTJ XXXV*. 141—51.

Pennanen, E. V. 1972. "Current views on WF". *Neuphilologische Mitteilungen 73*. 292—308.