A long and productive debate has been alive from the end of the Seventies up to now on theoretical and quantitative aspects of morphological productivity (e.g. Aronoff 1976, 1983; Booj 1977; Baayen 1992, 1993; Baayen, H., & A. Renouf. 1996; and, particularly on Italian language, Gaeta, L., & D. Ricca 2002, 2004), pointing out methodological and speculative issues related to affix usage in large corpora. On the other hand, while observing lexical roots, loan words usage and trends, we might wonder if it might be useful to introduce what analogously can be called lexical productivity?

When we try to describe and measure neologism’s behaviour in language, we cannot avoid observing, besides textual usage and diachronic trends of the single lexeme, the fact that some new roots tend occur often in other grammatical functions, as basis for affixation, as elements of compounds, and in specific collocations. For example, when dealing with loan words in Italian, such as internet, we detect also: free internet, internet banking, internet broker, internet café, internettiano, internettista, internettizzare, internettizzato, internavigatore, internauta, internavigatrice. The simple presence of such related usages make lexemes more integrated in the lexical system and in the speaker’s linguistic conscience.

The idea of monitoring and defining lexical productivity is far less common than it’s morphological counterpart. We find traces of qualitative and quantitative definitions in Kocourek (1996), Nakagawa (2000), Diaz Vera (2003), Kageura (2004) where simple reports of number of types, tokens and compounding aspects are accounted for. What is needed is a new measure able to report synchronic and diachronic productivity. Synchronic lexical productivity can be viewed as an index of textual usage in a given corpus, able to include as its parameters: a) number and frequency of different grammatical usages of the lexeme (e.g. download as a noun and as a verb), number and frequency of derivational lexemes (e.g. jeans > jeanseria; hobby > hobbista); c) number and frequency of compound words formed with the lexeme as an element (and in this case weighting roles as head of the compound and position) (e.g. mail > web-mail, voice-mail, junk-mail), including recursive classes of type a, b and c; d) number and frequency of collocations. Diachronic lexical productivity should be seen as indicating positive, negative or multimodal trends in lexical productivity observed at regular intervals (a slight modification of a measure presented in Bolasco 2005 will be proposed). It is possible to suggest the substitution of the mere frequency with usage index, only if segmentation for the determination of dispersion can be homogeneously applied to each diachronic interval.

Quantitative measures presented will be tested on a Italian corpus of five years of the newspaper “La Repubblica” (1996-2000, about 22 million word token each year) with approximately 110 million word occurrences. A number of recently attested loan words will be observed from the point of view of synchronic and diachronic productivity as well as in correlation with textual usage of the root forms analyzed. A further classification of loan words according to usage trends and productivity evidence will be proposed.

Theoretical questions related to the delicate phase of weighting different parameters that need to be included in the formulas, as well as questions about what kind of lexemes can profitably be described in term of lexical productivity (neologisms, loan words, etc.) will be addressed. Theoretical and methodological similarities and differences with morphological productivity issues will also be focused.
References