Lexical categories or frequency effects? A feedback from quantitative methods applied to psycholinguistic models in two studies on Italian.

Francesca Franzon[*], **Giorgio Arcara**[°], **Chiara Zanini**[*]
[*]Dipartimento di Neuroscienze, Università degli Studi di Padova
[°]IRCSS Ospedale san Camillo, Lido di Venezia
{francescafranzon7;giorgio.arcara}@gmail.com;
chiara.zanini.2@unipd.it

Abstract

English. We examined two issues concerning Italian Number morphology: the phenomena related to mass and count nouns and to the plural dominance. By taking into account quantitative data from corpora and subjective frequency ratings in three mixed effect models, we found that differences in participants' performance in two lexical decision tasks could be better captured as differences in frequency rather than in terms of effects of lexical categories.

Italiano. *In questo studio sono stati posti a confronto due fenomeni pertinenti alla morfologia nominale di Numero in italiano: la contabilità dei nomi e la dominanza plurale. Integrando i dati quantitativi provenienti dai corpora e da due studi di rating in un'analisi statistica condotta tramite modelli a effetti misti, risulta che le differenze nella prestazione dei partecipanti in due studi di decisione lessicale sono riconducibili a effetti di frequenza piuttosto che alla presenza di tratti lessicali categoriali.*

Introduction

The role of frequency in lexical retrieval is well known for what concerns psycholinguistic studies (since, at least, Forster & Chambers, 1973): the higher the frequency of a word, the faster its retrieval. Generally, the singular form of a noun is more frequent than the corresponding plural, and thus retrieved faster. However, some nouns (e.g. *stelle*, 'stars') do occur more frequently in the plural than in the singular: the phenomenon is known as plural dominance. Plural dominant nouns are in fact accessed faster in the plural form than in the singular. Since such nouns are not identifiable as a homogeneous group by means of some semantic features, the phenomenon has been explained as a mere effect of the frequency of occurrence of the forms (Baayen et al., 1996; 1997; 2007; Biedermann et al., 2013).

While the plural dominance seems to be unrelated to grammatical constraints, another phenomenon involving Number morphology seems to be grammatically grounded instead, namely the mass-count issue (Borer, 2005; Cheng, 1973; Chierchia, 2010; Jackendoff, 1991). Nouns referring to countable entities are called 'count nouns' (*anello*, 'ring'), nouns referring to uncountable entities are called 'mass nouns' (*burro*, 'butter'). Some constraints rule the possibility for the two types of nouns to occur in some morphosyntactic contexts, for example count nouns cannot occur in the singular after a quantifier (**molto anello*, 'much ring'), while mass nouns cannot occur with numerals or the indeterminate article (**un burro*, 'a butter'). For what concerns Number morphology, mass nouns should occur only inflected in the singular (but for a deeper discussion, see i.a. Acquaviva, 2013; Marcantonio & Pretto, 2001; Pelletier, 2012).

Previous lexical decision tasks have pointed out to some differences in the processing of count nouns with respect to mass nouns, which would require longer response times (RTs) (i.a. Mondini et al. 2009; Gillon et al. 1999). In the light of these results, it has been proposed that an additional lexical feature has to be computed for mass nouns as compared to count nouns.

While psycholinguistic studies on plural dominance have relied on relative frequency of singular and plural forms in the selection of stimuli and in results analysis, even the most recent experimental studies on the mass-count issue have not quantified the actual occurrence of the exper-

imental stimuli in mass context and in count context: nouns have rather been assigned to a mass or to a count category on the basis of the experimenters' judgments. Quantitative data on syntactic contexts can instead provide a better estimate of the frequency of use of nouns as countable or uncountable: in the present study we relied on the actual occurrence of nouns in the different syntactic contexts in assigning them to the "mass" or to the "count" experimental list.

We will describe and put into comparison two lexical decision tasks, concerning the phenomena of mass-count and of plural dominance respectively. We will explore the possibility that the mass-count effects described in psycholinguistic literature could be better explained in terms of frequency of occurrence, as it is recognized by most literature with respect to plural dominance. We hypothesize that the frequency of occurrence of the word form (inflected in the singular or in the plural) will predict the RTs in lexical decision tasks contrasting mass and count nouns, as well as in the ones concerning the plural dominance issue. The frequency of occurrence will be measured by means of two subjective frequency rating studies and in the corpus ItWaC (Baroni et al. 2009). We will rely on quantitative measures to categorize experimental stimuli. Measures of plural dominance of nouns will be based on the ratio between their occurrence in the plural and in the singular; the mass and count experimental nouns will be categorized considering their distribution with respect to mass and count morpho-syntactic contexts.

1 First study: mass and count nouns

1.1 Rating and corpus analysis

448 concrete nouns, namely 224 nouns inflected both in the singular and in the plural, were selected following the theoretical definitions given in traditional grammars. The list included the plural of 45 nouns for which only singular occurrences would be expected on a normative basis (pure "mass" nouns such as *burro* 'butter' - **burri* 'butters').

A questionnaire was designed in order to evaluate the subjective frequency of the 448 nouns following the methods used in previous literature (Ferrand et al., 2008). The questionnaire was administered online by means of the SurveyMonkey platform. 126 informants participated in this study (age: range = 22 - 76 years, mean = 36.2, SD = 12.46; years of education: range = 8-21). Participants were instructed not to express normative judgments, but to focus on the frequency they had heard or read the words; they had to assign a score to the frequency of the nouns on a 7-point Likert scale, ranging from 0 = "never heard or seen" to 6 = "more than once a day". The nouns in the questionnaires were presented to each participant in a different random order.

Score mean	Singular	Plural
$n = 0$	0	0
$0 < n \leq 1$	0	7
$1 \leq n \leq 2$	3	47
$2 \leq n \leq 3$	45	60
$3 \leq n \leq 4$	88	63
$4 \leq n \leq 5$	70	36
$n > 5$	14	7

Table 1: Distribution of the subjective frequency scores.

Absolute frequency of the aforementioned nouns was collected on the ItWaC corpus (Baroni et al., 2009). A positive correlation was found between corpus frequency and subjective frequency: $r(446) = 0.75$, $p < .001$. In order to disambiguate the mass use from the count use of the nouns presented in the rating questionnaire, we designed queries in CQP syntax following the methods described by Katz & Zamparelli (2012). The occurrence of nouns with determiners such as the indeterminate article and quantifiers were used to trace the occurrence in unambiguous count or mass context.

1.2 Lexical decision task

From the initial list of 224 nouns, 80 nouns were selected and presented both in the singular and in the plural (totally 160 experimental stimuli). These stimuli were selected to span as uniformly as possible across the range of possible values of subjective frequency in order to use the subjective frequency as a continuous variable in the analysis. From the 80 nouns we classified as "mass" the 18 top mass-used nouns with the highest mass frequencies and values of count frequencies that were not among the top 18; we classified as "count" the 18 top count-used nouns with the highest count frequencies and values of mass frequencies that were not among the top 18. The nouns were presented both in the singular and in the plural (totally 72). The remaining stimuli were not categorised in such terms. Experimental stimuli are displayed in table 2. The

final list included 240 filler words, consisting in 80 adjectives and 160 phonotactically plausible non-words.

	N. of items	Corpus Frequency	Subjective Frequency	Length
All stimuli	160	11850.32 (27239.65)	3.29 (1.18)	6.41 (1.66)
"Mass" nouns: singular	18	26204.88 (28831.43)	4.36 (0.57)	6.22 (1.89)
"Mass" nouns: plural	18	824 (1187.38)	1.95 (0.72)	6.28 (1.96)
"Count" nouns: singular	18	38570.05 (54194.95)	4.09 (0.84)	5.78 (1.31)
"Count"nouns: plural	18	24365 (36455)	4.07 (0.80)	5.89 (1.27)

Table 2: Psycholinguistic properties of experimental stimuli.

60 Italian native speakers participated in the experiment (mean age = 23.5, SD = 2.37; years of education: mean = 15.16, SD = 1.64). Participants saw a series of letter strings presented at the center of the screen one at a time. They had to press a key if they thought the string was an Italian word, another key in the converse case.

1.3 Results

Results were analyzed by means of mixed effect models (Baayen, Davidson & Bates, 2008). In the model 1, summarized in table 3, we included the 72 stimuli classified as mass and count nouns. We considered as predictors: *category (mass/count)*, *Number (singular/plural)*, *corpus frequency*, *subjective frequency* and *orthographic length*. Results show significant effects of length (longer RTs for longer items), of Number (longer RTs for plurals) and of subjective frequency (longer RTs for low subjective frequency).

Fixed effect	Coefficient	Standard Error	df	t	p-value
Intercept	6.56	0.05	95.18	130.53	<0.001
Number= plural	0.37	0.02	64.33	2.04	0.04
Subjective frequency	-0.04	0.007	74.09	-4.27	<0.001
Orthographic length	0.009	0.004	65.86	2.077	0.04

Table 3: Results of model 1.

In model 2, summarized in table 4, we included all the 160 stimuli. We considered as predictors: *Number* (singular/plural), *corpus frequency*, *subjective frequency* and *orthographic length*. Results show significant effects of length (longer RTs for longer items), of corpus frequency (longer RTs for low corpus frequency) and of subjective frequency (longer RTs for low subjective frequency).

Notably, the predictor c*ategory* is not significant (p = 0.85); *corpus frequency* is a significant predictor in model 2 (p = 0.03), but it only approached significance in model 1 (p = 0.05). Possibly, in model 1 *Number* is a significant predictor because the categorised items represent a subset that differ for frequency of occurrence in the plural. In fact, in model 2, in which both categorised and not categorised items were considered, no effect of Number was found.

Fixed effect	Coefficient	Standard Error	df	t	p-value
Intercept	6.73	0.04	219.42	172.38	<0.001
Corpus frequency	-0.009	0.004	155.55	-2.16	0.03
Subjective frequency	-0.05	0.008	152.19	-5.37	< 0.001
Orthographic length	0.008	0.004	2.47	2.11	0.04

Table 4: Results of model 2.

2 Second study: plural dominance

2.1 Rating and corpus analysis

The ItWaC corpus was queried to obtain the frequency of occurrence of the singular and the plurals of nouns displaying the most common inflectional patterns (-o/-i; -a/-e). We discarded from testing material compounds, derived nouns and the nouns that differ for orthographic length or phonological form between singular and plural (e.g. *occhio* - *occhi* 'eye –eyes'). The remaining nouns were then ordered on the base of their plural dominance defined as the ratio plural frequency/singular frequency. We calculated stem frequency of nouns and selected 284 nouns uniformly span across the range of possible values of frequency.

A questionnaire was created in order to test the subjective frequency of the 284 selected nouns, both in the singular ad the plural (568 experimental items). The questionnaire was administered following the same methods described previously (§2.1). 150 Italian native speakers participated in the study (age: range = 18 – 69, mean = 29; years of education: range = 8-21). The distribution of the subjective frequency is plotted in Table 5. A positive correlation was found between the singular and plural forms

of nouns within the corpus ($r(282) = 0.70$, $p < .001$) and within the rating ($r(282) = 0.91$, $p < .001$).

Score mean	Singular	Plural
n = 0	0	0
0 < n ≤ 1	0	1
1 ≤ n ≤ 2	19	20
2 ≤ n ≤ 3	88	100
3 ≤ n ≤ 4	139	131
4 ≤ n ≤ 5	31	27
n > 5	7	5

Table 5: Distribution of the subjective frequency scores.

2.2 Lexical decision task

A lexical decision study was carried out, following the same methods described in §2.2. From the 284 nouns mentioned in §3.1, we chose: the 30 nouns with the highest ratio of plural dominance, the 30 nouns with the lowest ratio of plural dominance, the 30 nouns whose ratio between singular ad plural was the closest to 1 (see table 6). Each noun was presented in the singular and in the plural (totally 180 experimental stimuli). The final list included 364 filler words, consisting in 184 adjectives and 180 phonotactically plausible non-words.

43 Italian native speakers participated in the experiment.

Dominance (mean Pl/Sg)	Morphological Number	N. of items	Corpus Frequency	Subjective Frequency	Orthographic Length
Plural (3.61)	Singular	30	5260.3 (7547.43)	3.31 (0.77)	6.33 (1.09)
	Plural	30	19026.46 (25558.41)	3.48 (0.79)	
Singular (0.16)	Singular	30	25596.9 (44944.15)	3.44 (0.91)	6.13 (1.13)
	Plural	30	4276.3 (7186.03)	3.23 (0.79)	
Equal (0.9)	Singular	30	35430.33 (99471.4)	3.13 (0.57)	6.16 (1.17)
	Plural	30	31921.7 (93584.35)	3.1 (0.59)	

Table 6: Psycholinguistic properties of experimental stimuli.

2.3 Results

Results were analysed by means of mixed effect models (Baayen, Davidson & Bates 2008). In model 3, summarized in table 7, we considered as predictors: *category (plural/singular/equal dominant)*, *Number (singular/plural)*, *corpus frequency*, *subjective frequency* and *orthographic length*. Results show significant effects of length (longer RTs for longer items), of corpus frequency (longer RTs for low corpus frequency) and of subjective frequency (longer RTs for low subjective frequency).

Fixed effect	Coefficient	Standard Error	df	t	p-value
Intercept	6.79	0.04	211.64	137.79	<0.001
Corpus frequency	-0.02	0.003	171.55	-7.23	<0.001
Subjective frequency	-0.03	0.007	170.17	-4.48	< 0.001
Orthographic length	0.009	0.004	165.93	2.03	0.04

Table 7: Results of model 3.

3 Discussion and conclusions

In this study we applied quantitative methods in the selection of experimental stimuli used in the two lexical decision tasks. In both tasks, results from the three models showed effects of subjective frequency and corpus frequency but not of category in written word recognition. For what concerns the plural dominance issue, this result was in line with previous literature. For what concerns the mass-count issue, our results are unexpected instead. Remind that frequency of occurrence in mass and count contexts was used to avoid biases in categorization of stimuli. Nevertheless, we did not observe differences in RTs between the two so categorized groups of nouns. Thus, we suggest that there is no need to postulate the computation of a lexical feature related to countability or uncountability in nouns. We propose that the fact that a noun is considered "mass" is better described as an epiphenomenon of the distribution of noun with respect of syntactic contexts. However the possibility for a noun to occur in the different syntactic contexts does not predict lexical decision RTs: frequency, as measured in the corpus and by the rating study, is the predictor of the lexical access times with respect to words presented in isolation. In this sense, the mass-count issue is similar to the plural dominance phenomenon: even in that case, there is no need to assume the presence of a feature marking plurality, as the frequency of the inflected form is sufficient to account for the observed effects in lexical decision tasks.

The frequency of occurrence of nouns considered as a continuous variable is a better predictor of RTs than a distinction attributed to alleged lexical categories both in the case of phenomena

seemingly unrelated to core grammar rules, like the plural dominance, as well as in phenomena that have traditionally been described as grammar based, like the mass-count issue.

References

Acquaviva, P. (2013). *Il nome*. Roma: Carocci.

Baayen, H., Burani, C., & Schreuder, R. (1996). Effects of semantic markedness in the processing of regular nominal singulars and plurals in Italian. *Yearbook of morphology*, Springer Netherlands, 13-33.

Baayen, R. H., Dijkstra, T., & Schreuder, R. (1997). Singulars and plurals in Dutch: Evidence for a parallel dual-route model. *Journal of Memory and Language*, 37(1), 94-117.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390-412.

Baayen, R., Levelt, W., Schreuder, R., & Ernestus, M. (2007). Paradigmatic structure in speech production. *Proceedings from the Annual Meeting of the Chicago Linguistic Society*, 43(1): 1-29. Chicago Linguistic Society.

Balota, D. A., Pilotti, M., & Cortese, M. J. (2001). Subjective frequency estimates for 2,938 monosyllabic words. *Memory & Cognition* 29(4), 639-647.

Baroni, M., Bernardini, S., Ferraresi, A., & Zanchetta, E. (2009). The WaCky Wide Web: A Collection of Very Large Linguistically Processed Web-Crawled Corpora. *Language Resources and Evaluation* 43 (3), 209-226.

Biedermann, B., Beyersmann, E., Mason, C., & Nickels, L. (2013). Does plural dominance play a role in spoken picture naming? A comparison of unimpaired and impaired speakers. *Journal of Neurolinguistics*, 26(6), 712-736.

Borer, H. (2005). *In name only*. Oxford: OUP.

Cheng, C.-Y. (1973). Response to Moravcsik. J. Hintikka, J.M.E. Moravcsik, & P. Suppes (eds.). *Approaches to Natural Language*. Dordrecht: Reidel, 286-288.

Chierchia, G. (2010). Mass nouns, vagueness and semantic variation. *Synthèses* 174, 99-149.

Ferrand, L., Bonin, P., Méot, A., Augustinova, M., New, B., Pallier, C., & Brysbaert, M. (2008). Age-of-acquisition and subjective frequency estimates for all generally known monosyllabic French words and their relation with other psycholinguistic variables. *Behavior Research Methods* 40 (4), 1049-1054.

Forster, K. I., & Chambers, S. M. (1973). Lexical access and naming time.*Journal of verbal learning and verbal behavior*, *12*(6), 627-635.

Gillon, B., Kehayia, E., & Taler, V. (1999). The mass/count distinction: Evidence from on-line psycholinguistic performance. *Brain and Language* 68, 205-211.

Jackendoff, R. (1991). Parts and boundaries. *Cognition* 41, 9-45.

Katz, G. & Zamparelli, R. (2012). Quantifying Count/Mass Elasticity. Choi, J. et al. (eds). *Proceedings of the 29th West Coast Conference on Formal Linguistics*. Somerville, MA: Cascadilla Proceedings Project, 371-379.

Kulkarni, R., Rothstein, S., & Treves, A. (2013). A Statistical Investigation into the Cross-Linguistic Distribution of Mass and Count Nouns: Morphosyntactic and Semantic Perspectives. *Biolinguistics* 7, 132-168.

Kuperman, V., & Van Dyke, J. A. (2013). Reassessing word frequency as a determinant of word recognition for skilled and unskilled readers. *Journal of Experimental Psychology: Human Perception and Performance* 39(3), 802.

Marcantonio, A. & Pretto, A. M. (2001). Il nome. L. Renzi, G. Salvi, & A. Cardinaletti (eds.). *Grande grammatica italiana di consultazione*. Bologna: Il Mulino, 329-346.

Mondini, S., Kehaya, E., Gillon, B., Arcara, G., & Jarema, G. (2009). Lexical access of mass and count nouns. How word recognition reaction times correlate with lexical and morphosyntactic processing. *The Mental Lexicon* 4, 354-379.

Pelletier, F. J. (2012a). Lexical Nouns are Neither Mass nor Count, but they are Both Mass and Count. D. Massam (ed.). *A Cross-Linguistic Exploration of the Count-Mass Distinction*. Oxford: OUP, 9-26.

Williams, R., & Morris, R. (2004). Eye movements, word familiarity, and vocabulary acquisition. *European Journal of Cognitive Psychology* 16(1/2), 312–339.