Bilingual Frequency in a Favorable Context (BFFC) in the Italian dialectal area. A protocol for the analysis of geminate lateral retroflexion in Antona (MS)

From the end of the past century, usage-based approaches to the analysis of language variation have increasingly become widespread in experimental linguistic research. In particular, exemplar theory posits the storage of detailed eventive traces in the mind of the individual, comprehensive of linguistic, social and environmental features. In this framework, the structure of the mental clouds (e.g. [1]) of exemplars is of central importance in determining the individual choices, revolving around the properties of salience, recency and density, i.e. frequency of exposure to similar events. One of the main question scholars had to answer in order to test the effectiveness of this model is what kind of frequency exemplar theory is concerned with. Bybee [2] pointed out that researchers should not consider the raw frequency of a word as the main predictor of its role in a variation phenomenon; instead, since the linguistic components of an exemplar are purely phonetic, the frequency of a word in across-boundaries contexts of segmental variation is of greater concern for usage-based models. This idea remained largely untested until Raymond and Brown's proposal for a quantification method for this parameter, named *Frequency in a Favorable Context* (FFC) and calculated as the frequency of a word in a variating context divided by its overall frequency [3]. Another related issue refers to the role of individual frequency of use in shaping exemplar clouds. From the moment that each linguistic lifespan is indeed a «personal odyssey» ([4]: 19) of events, the effects of individual patterns should be proved at least comparable to those derived from a corpus. Drager [5] helped clarifying this point, integrating months of ethnographic fieldwork with experimental phonetic inquiries. In her data, individual frequencies of occurrence of *like*, shaped by its context and grammatical function, had a significant role in triggering sociophonetic variation.

In 2015, Brown [6] expanded her line of studies to bilingual speakers. The scholar examined a corpus of oral interviews in a bilingual (Spanish-English) community in New Mexico, searching for the effects of word frequency on Spanish /d/ weakening, mainly occurring in post-vocalic contexts. She thus revised the FFC formula proposing the *Bilingual Frequency in a Favorable Context* (BFFC), apt to verify the weight of existing cognate (i.e. phonologically and semantically similar) English forms on the Spanish variation phenomenon at hand. In the BFFC, the corpus frequency of the English cognate is integrated as a non-variating context and thus added to the overall Spanish frequency of the word. The experiment proved that the English cognates had a strong impact in inhibiting the inquired variation, theoretically confirming that bilingual speakers cluster all their communicative events in a single exemplar cloud.

In this paper I argue that this last point is crucial for the adjustment of sociophonetic methodologies to the Italian linguistic landscape. Italian dialects are in fact «'sisters' of Italian, locally divergent developments of the Latin originally spoken in Italy» ([7]: 2); from a sociolinguistic point of view, it has been noted that the structural distance between the dialects and the standard language «justify treating the Italian situation as bilingual rather than merely bidialectal» ([8]: 394). The probability of the emergence of local traits in the dialectal practice would thus be conditioned by the BFFC, with a negative influence exerted by Standard Italian cognates. The main hurdle to overcome in order to test the BFFC in an Italian setting is the impossibility to extract dialectal frequencies from an adequate corpus. For this reason, I suggest using previously retrieved subjective frequency estimates for the dialectal component of the formula, in line with the above discussed compatibility between individual and corpus-grounded values. Subjective estimates have been proved to be closely correlated to objective, corpus-based frequencies, both from the theoretical [9] and experimental [10] points of view; moreover, reliable measures can be built from a reasonably small percentage [10; 11] of the represented population. The standard protocol for collecting subjective estimates has been established by Balota and colleagues [11]: interviewees are simply asked to rate on a Likert scale the frequency of their encounters with specific words, ranging from 1=*never* to 7=*several times a day*.

After a review of the relevant literature, this paper focuses on the steps of a protocol from a research in progress on geminate lateral retroflexion (traditionally transcribed as *-ll-* > *-ḍḍ-*) in Antona (MS). This feature, first described by Bottiglioni [12], covers the area of the Apuan Alps, not without internal variability [13]. In Antona, a small mountain village in the Frigido valley above the chief-town of Massa, lateral retroflexion is today the flagship feature of local linguistic identity, in a general picture of gradual loss of dialectal competence and consequent advergence ([14]: 82) of the dialect to Standard Italian. For the sake of simplicity, in this explorative study the phenomenon will be observed only in word-internal positions, excluding syntactic doubling and equating the dialectal FFC value to the overall frequency of the word. The protocol consists in: a) extraction of singular nouns, ordered by frequency of occurrence and containing geminate laterals, from an Italian corpus; b) submission of the list to a small number of native informants, in order to exclude the Italian nouns that do not share the cognate status with their dialectal counterparts; c) submission

of the derived list of nouns to a socially distributed group of native informants to assess subjective frequency estimates, revising Balota and colleagues' instruction with the indication of the locality (i.e. Antona) of the encounters with the noun; d) calculation of the BFFC as the mean of the subjective estimate ratings divided by the same value plus the log. objective frequency of the Italian cognate; d) selection of a subset of nouns, equally distributed by their BFFC ratings; e) pairing of these nouns with graphic representation of their referents, using a corpus of images built for psycholinguistic testing (e.g. [15]); f) picture naming task submitted to a socially distributed group of native interviewees, specifying the informal communicative context of the simulation. If BFFC has an impact on the emergence of dialectal variants, a correlation between its values and the number of elicitations of lateral retroflexion in the task is expected. At present time, a list of 185 singular nouns containing a geminate lateral has been extracted from *ItWac* [16] through a case-insensitive query. The list has been presented to three informants that have cancelled the non-cognate nouns, writing down their dialectal equivalents (e.g. It. *pipistrello* ≠ Ant. *parpagon*; see [17]: 378). Given the young age of these informants, this phase of the research still needs more feedbacks before moving to the subjective frequency estimates. In fact, we could expect that older speakers retain a wider pool of dialectal lexemes, adding depth to our description of the interaction between the two linguistic systems in Antona.

Bibliography

[1] Pierrehumbert, J. B. (2001). Exemplar dynamics: Word frequency, lenition, and contrast. In J. Bybee and P. Hopper (eds), *Frequency effects and the emergence of lexical structure*. John Benjamins: Amsterdam, 137-157.

[2] Bybee, J. L. (2002). Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. In *Language Variation and Change*, 14, 261–90.

[3] Raymond, W.D. and Brown, E.L. (2012). Are effects of word frequency effects of context of use? An analysis of initial fricative reduction in Spanish. In S. Th. Gries and D. S. Divjak (Eds.), *Frequency Effects in Language. Vol 2: Learning and Processing*. Mouton de Gruyter: The Hague, 35-52.

[4] Foulkes, P. (2010). Exploring social-indexical knowledge: A long past but a short history. In *Laboratory Phonology*, 1(1), 5-39.

[5] Drager, K. (2011). Sociophonetic variation and the lemma. In *Journal of Phonetics*, 39, 694-707.

[6] Brown, E.L. (2015). The role of discourse context frequency in phonological variation: A usage-based approach to bilingual speech production. In *International Journal of Bilingualism*, 19(4), 387-406.

[7] Maiden, M. and Parry, M. (eds, 1997). *The Dialects of Italy*. Routledge: London and New York.

[8] Berruto G. (1997). Code-switching and code-mixing. In M. Maiden and M. Parry (eds.), *The Dialects of Italy*. Routledge: London and New York, 394-400.

[9] McGee, I. (2008). Word Frequency Estimates Revisited—A Response to Alderson (2007). In *Applied Linguistis*, 29(3), 509–514.

[10] Thompson, G. L. and Desrochers, A. (2009). Corroborating biased indicators: global and local agreement among objective and subjective estimates of printed word frequency. In *Behavior Research Methods*, 41(2), 452-471.

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

[12] Bottiglioni, G. (1911). Dalla Magra al Frigido. Saggio fonetico. In *Revue de Dialectologie Romane*, 3, 77-143.

[13] Savoia, L. (1980). Fonologia delle varietà apuane e garfagnine: consonantismo. In *Studi Urbinati di Storia, Filosofia e Letteratura*, II, 233-293.

[14] Berruto, G. (2005). Dialect/standard convergence, mixing, and models of language contact: the case of Italy. In P. Auer, F. Hinskens and P. Kerswill (eds), *Dialect Change. Convergence and Divergence in European Languages*. University Press: Cambridge, 81-95.

[15] Szekely, A., Jacobsen, T., D'Amico, S. *et al.* (2004), A new on-line resource for psycholinguistic studies. In *Journal of Memory and Language*, 51(2), 247-250.

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

[17] Mosti, E. and Nancesi, M. (2005). *Dizionario enciclopedico del vero dialetto massese*, CRD: Massa.