

## Phonetic Variation and Syllabic Structures in Italian and German Speech

Loredana Schettino<sup>1,2</sup>, Maria Di Maro<sup>1</sup>, Antonio Origlia<sup>1</sup>, Francesco Cutugno<sup>1</sup>

<sup>1</sup>*University of Naples Federico II, Italy;* <sup>2</sup>*University of Salerno, Italy*

In human spontaneous speech, the phonetic realization of words may vary and deviate from their “canonical form”. These deviations may concern segment changes, such as vowel centralization or consonant lenition, or segment deletion [1]. Hence, they have been usually referred to as “reduction phenomena”. However, numerous studies in different languages have shown that these phenomena are not the exception to the rule but rather represent the rule in spoken language communication [see 2], which may be due to linguistic factors, like prosodic features, lexical category, discursive function, and be affected by the sociolinguistic factors related to the communicative situation [3]. Therefore, studies investigating these phenomena are required to contribute to unveiling the functioning of the speech production and comprehension mechanisms, improve ASR performances and support foreign language learning.

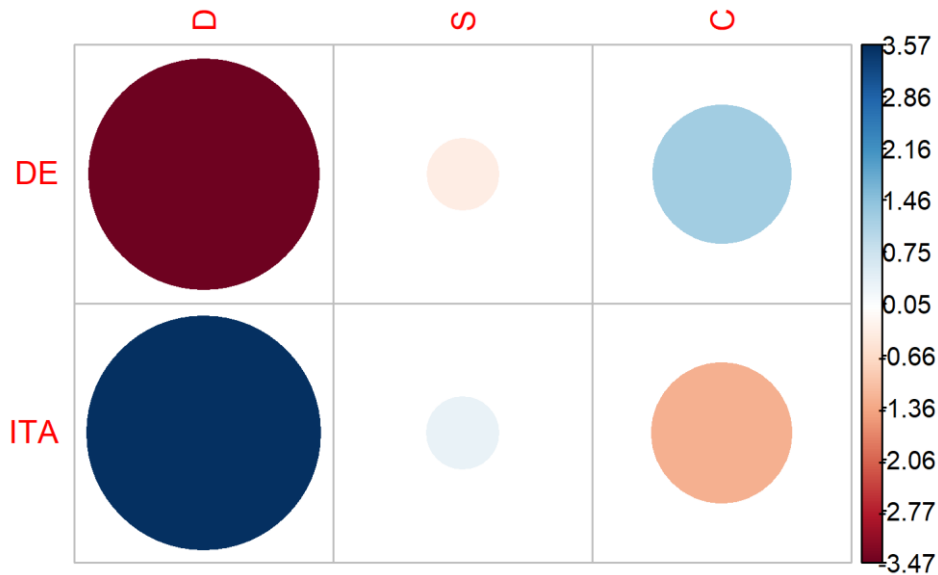
This study aims at exploring the phonetic variation (and reduction) patterns that may be observed in the speech chain, and their relation to linguistic structures, in particular syllabic structures and stress, considering two different languages: Italian and German.

The investigation was based on the Italian and German datasets of the Nocando corpus [4] which consist of spoken narrative texts mostly produced by university students (11 Italian and 8 German), so the communicative situation was the same for both datasets. The audio files and transcriptions were processed using the WebMAUS Basic services [5]. The phonological and phonetic transcriptions were manually edited in Praat [6] and syllabified according to the principles of onset maximization and sonority sequencing [7]. The alignment between the sequence of phonological syllables and the one of phonetic syllables was evaluated using SCLITE, which is a tool included in the Speech Recognition Scoring Toolkit (SCTK) provided by the NIST. The evaluation output reports as “Deletion” the cases where a phonetic syllable is found in place of two phonological ones; as “Substitutions” the cases where a phonetic syllable is different from the corresponding phonological one; as “Correct” the cases where the phonetic and the phonologic syllables are alike. The phonological syllables were also annotated for their phonetic structure and saliency (lexical stress). For statistical analysis, Generalized Linear Mixed Models were fitted: the evaluation levels were considered as dependent variables; the Languages, Syllabic Structure and Lexical Stress as interacting independent variables, and Speaker as random effect.

The analysis is based on 2050 phonological syllables, that is 950 in the Italian and 1100 in the German dataset. Significant results reveal that the language, the syllabic structure and its saliency all can represent significant predictors of phonetic variation. More specifically, Italian data present significantly more Deletion phenomena than German data (see Figure 1). CV represents the most stable structure, which is significantly less subjected to variation, in both Italian and German. Then, as expected, unstressed syllables are significant predictors of Deletion phenomena, however, only for German they also significantly predict Substitution phenomena.

Further investigation is concerning the qualitative analysis of Deletion and Substitution phenomena.

Figure 1. Correlation matrix plot between the evaluation output (D=Deletion, S=Substitution, C=Correct) and the languages (DE=German, ITA=Italian).



- [1] Adda-Decker, M., & Snoeren, N. D. Quantifying temporal speech reduction in French using forced speech alignment. *Journal of Phonetics*, 39(3), 261-270, 2011.
- [2] Cangemi, F., & Niebuhr, O. Rethinking reduction and canonical forms. *Rethinking Reduction*. De Gruyter Mouton, 277-302, 2018.
- [3] Savy, R. & Cutugno, F. Ipoarticolazione, riduzione vocalica, centralizzazione: come interagiscono nella variazione diafasica? In Cutugno f. (ed), *Fonetica e fonologia degli stili dell'italiano parlato*. Atti delle VII Giornate di Studio del Gruppo di Fonetica Sperimentale 24, 177-194, 1997.
- [4] Brunetti, L., S. Bott, J. Costa and E. Vallduví. 'A multilingual annotated corpus for the study of Information Structure', In Konopka et al (eds), *Grammar & Corpora 2009*, 3rd international Conference, Mannheim, 22-24 sept. 2009, Gunter Narr Verlag, 2011.
- [5] Kisler, T. and Reichel U. D. and Schiel, F. Multilingual processing of speech via web services, *Computer Speech & Language*, Volume 45, September 2017, pages 326–347, 2017.
- [6] Boersma, P., & Weenink, D. Praat: doing phonetics by computer. *Computer program*, 1992-2022.
- [7] Nespor, M. A. *Fonologia*. Il Mulino, 1993.