NON-NATIVE PHONOLOGY: ITALIAN VOWEL SYSTEM OF GERMAN SPEAKERS

VALENTINA SCHETTINO
University of Naples "L'Orientale"

Abstract

The present study investigates the production of Italian vowels by German speakers, both from a phonological and from an acoustic perspective. Aim of the work is to check possible transfer from the native language, to examine the quality and quantity of this transfer and to interpret the results in the framework of Natural Phonology (Donegan & Stampe, 1979). Our corpus consists of two German native speakers, recorded both in Italian and in German. Particular attention is devoted to the examination of formant values: expected and actually recorded mean frequencies are compared; in this respect, deviations from expected values are identified and interpreted as transfer processes: in some case transfer comes from the L1, but we individuate also transfer phenomena from the L2. Additional results about contextual application of transfer processes and possible explanations are discussed. It is concluded that context and variability are crucial in spontaneous speech, and that a good phonological theory should highlight the naturalness of such phenomena.

Introduction

This paper aims at confronting Italian and German vowel systems; specifically, we intend to observe the articulatory behaviour of German native speakers living in an Italian-speaking environment–in the city of Naples. We focused on the study of vowels in order to avoid superficial analyses: in fact, an examination of the whole phonological chart in both language could have been hazardous due to insufficient data. On the contrary, we were able to shed some lights on L2 vowel production – as will be shown in the next paragraphs – despite scarcity of materials.

The theoretical framework adopted in this work is that of Natural Phonology, a functionalist approach developed by the American linguists David Stampe and Patricia Donegan in the late sixties and seventies[1]. The main guideline of Natural Phonology is of articulatory nature: namely, every phonological element reflects the fact that language–trivially–is spoken; every communication act, then, expresses a precise intention of the speaker and a consequent reaction of the listener. The latter correlation shows that the language is nothing but a reflection of human reality:

> This is a natural theory, in the sense established by Plato in the *Cratylus*, in that it presents language (specifically the phonological aspect of language) as a natural reflection of the needs, capacities, and world of its users, rather than a merely conventional institution.
>
> (Donegan & Stampe 1979:127)

According to Donegan & Stampe (1979), the phonology of a given language has to be interpreted as a set of innate phonological *processes*[2], which are revised and improved

[1] Natural Phonology was first introduced by David Stampe in his doctoral Thesis (1969) and then developed in Donegan & Stampe (1979).

[2] Processes are defined in Donegan & Stampe (1979:126) as "mental substitutions which systematically but subconsciously adapt our phonological intentions to our phonetic capacities".

through the experience of speech. In other words, processes result from the mutual relationship between phonetic skills and phonological restriction of those abilities. It should be stressed, however, that a natural process is first of all a mental action, based on phonological properties, not only on phonetic aspects, even though there is a clear articulatory basis.

Nevertheless, not every phonological alternation is classified as process. For the most conventional cases, in fact, Stampe introduces the concept of *rule*[3]: while processes are speaker-dependent, being restrictions on the language made by each particular speaker, rules are all the restrictions required by every speaker for spontaneous speech.

We do not have space here to appropriately refer to the whole phonological theory presented by Stampe and Donegan; Natural Phonology, however, has to be seen as a new, truly empirical approach in the field of linguistics. Other theoretical frameworks, in fact, appear to underestimate the role of articulation and spontaneous speech. For Chomsky and the generative group, for example, "an explanatory theory is one which provides […] a procedure for selecting the correct grammar for given data" (Donegan & Stampe, 1979:128). Structuralism, on the other side, focuses on distinctive features, avoiding further examination about the actual use of discrete units: "words are not only distinguished by sounds, they are made up of them. It is no less important that the sounds that constitute words be distinguishable than that they be pronounceable, combinable, and perceivable (articulate, audible)" (Donegan & Stampe, 1979:129).

Some linguists have objected that Natural Phonology is too obvious to be proved wrong (Kiparsky & Menn, 1977), others (Tse, 1980) have pointed at the evolution of comprehensive psycholinguistic studies that might contradict it. In any case, this theory has managed to put the attention back on spontaneous speech processes and real use of spoken language: for this reason, we will base the following empirical work on Natural Phonology.

1. Materials

Natural languages are characterized by a considerable amount of variability, which is clearly emphasised in spoken contexts. This variable nature is even more accentuated when it comes to forms of spontaneous speech, which is dramatically unpredictable from a scientific–acoustic point of view.

As a consequence, an efficient strategy is necessary, for example regarding the used tools or the corpus collection methodology, with a special focus on the linguistic aspects we intend to dwell on, in our case vowel phones. Therefore, the preparation of the material had to include the choice of mechanical equipment (microphones and recording programs), the study of the interviewing methodology, the selection of the segmentation and data analysis programs to create sonograms, spectrograms and the study of formants.

The choice of the speakers cannot be underestimated, too, as German dialects can be consistently different with respect to vowels (cf. Wiesinger, 1971; Löffler, 2003); for this reason, we interviewed two German-speaking subjects born in geographically opposite regions: the first comes from Rhineland, the second is Austrian, in order to provide a sufficiently differentiated dataset (see Figure 1).

[3] Rule are presented in Donegan & Stampe (1979:143) as a reverse of processes: "The principles which underlie alternations which are not process-governed […] we refer to as phonological rules". However, a clear, free-standing definition is not provided.

Both speakers, however, lived in Italy at the moment of the data collection, precisely in Naples. They both worked at the University of Naples "Federico II".

Figure 1 Geographical provenance of the speakers

Indeed, the interviews were conducted in an anechoic chamber at the Language Laboratory of the above-mentioned university "Federico II". For the recording phase we chose *Goldwave*[4], a very sophisticated and highly precise open source program. The files were recorded at 22050 Hertz, the best frequency for sound quality; standard headphones equipped with a microphone were used.

The computer software chosen for the analytical study of the data is another open source program called *Wavesufer*[5], widely used for phonetic studies and suitable both for the production of spectrometers and sonograms and for the acoustic analysis of the phonemes with relative formants; moreover, it provides the user with frequency and duration values. Every physical-acoustic detail of each registered segment is computed in a few seconds, with excellent results from the qualitative point of view. Furthermore, this program enables the simultaneous analysis of all variables.

1.1. The corpus

As regards the corpus, two German speakers were recorded: they both had been living in Italy for more than a year at the moment of the interview, but they came from geographically opposed German-speaking communities (cf. Figure 1). The whole recording process consisted of two different steps: in the first phase, Italian semi-spontaneous speech was elicited through map-task techniques[6]: speakers were asked to analyze two very similar pictures[7] and talk in their variety of Italian about the small differences in the illustrations[8]. This technique has been often applied for the registration of corpora of spoken dialogues (Albano Leoni et al., 2005), because it has several advantages, both for the interviewee and for the interviewer: speakers are put at ease by an informal setting and affordable tasks, in which they must cooperate. Moreover, it allows the achievement of the interviewer's stated aim ensuring a very high quality of the data: this type of interview, indeed, is highly

[4] Cf. Craig (2008), online at https://www.goldwave.com/.
[5] Cf. Sjölander & Beskow (2000), online at http://www.speech.kth.se/wavesurfer/.
[6] This technique was originally conceived by Brown et al. (1984) and then developed at Edinburgh for the acquisition of the corpus HCRC Map Task (Anderson et al, 1991).
[7] The pictures consisted of two comics of ducks at sea with a life jacket and a diver's mask; little differences appeared quite clearly in some details of the picture, for example the orientation of the beak or the colour of the life vest.
[8] *Spot the difference* pictures were borrowed from the Italian project CLIPS (en. Corpora and Lexycon of written and spoken Italian, Albano Leoni et al., 2005).

recommended (Cerrato, 2006) because it is suitable for recording a satisfactory set of phones. In fact, the objects in the pictures can be chosen according to the phonemes contained in the reference lexemes, obtaining a perfect set of produced vowels for the following analysis phase. There might be a chance that speakers do to produce the desired items: however, the choice of focusing on vowel only, together with the high probability of describing the differences in the expected way, helped us in the collection task.

The second phase of the registration consisted of a prepared set of questions in German, during which the two speakers acted first as interviewer and then as interviewee: this change of roles was required in order to avoid differences in the spontaneity level and in the amount of sentences actually registered by the two speakers.

In our opinion, the different methodologies used to elicit German and Italian data do not represent a drawback for the comparability of our results: indeed, in both cases we obtained semi-spontaneous productions, with prearranged semantic field and induced lexemes production. It is not relevant, then, if their answers are prompted by a prepared question or by a prepared image. The sentences produced by the speakers when acting as interviewers were not taken into consideration for the final analysis: in those sentences, in fact, we do not have semi-spontaneous speech, but read speech instead. Therefore, formants and other acoustic values may be quite different from those typical of spontaneous productions (Nakamura et al., 2008).

In both phases, the speakers were not able to see each other, in order to eliminate the risk of too extensive non-verbal communication and to maximize the dialogic interaction; they were also asked to avoid background noise as much as possible, such as coughing, chair movements, crackling etc. The simultaneous presence of both speakers has to be underlined as a positive value: in this way, in fact, environmental conditions are not altered and the objectivity of the data is preserved.

2. Analysis and results

With respect to the analysis, mean formants value of vowels were examined, with reference to previous Italian and German studies (Albano Leoni & Maturi, 2010; Sendlmeier & Seebode, 2006): all Italian vowels produced by German speakers (except [a]) presented values of F2 out of the norm- higher for back vowels and lower for front vowels. The complete results are shown in Table 1.

	F1 actual	**F1 expected**	**F2 actual**	**F2 expected**
[u]	268	**305**±55	920	**861**±135
[o]	303	**409**±58	1161	**1001**±257
[ɔ]	522	**554**±65	1214	**1055**±191
[a]	701	**708**±87	1454	**1466**±109
[ɛ]	490	**500**±77	1599	**1844**±181
[e]	283	**375**±63	1687	**2028**±195
[i]	283	**275**±61	1826	**2240**±160

Table 1: Mean formant values of Italian vowels produced by German native speakers

In particular, [o] and [e] show uncommon F1 values, too. In both cases, we could find a phonological motivation: as regards the former case, as the open-mid back vowel does not

exist in German in the lengthened form [ɔ:], many Italian words including this vowels are pronounced with a [o:] instead; close vowels, however, are articulated higher (and longer) in German: for this reason, F1 values – which are directly proportional to the openness degree and inversely proportional to the height – are lower. In the latter case, F2 values indicate lower degree of frontness and, as a consequence, a more central position in the vowel chart; this happens because of a transfer from the L1, in which [e] is always long and consequently closer and less central.

German vowels, on the contrary, are consistently produced as we expect them; full results can be observed in Table 2. As it can be seen, F1 and F2 values do not deviate from expectations more that 150 Hz, with the exception of F2 values of [ə].

	[i]	**[y]**	**[e]**	**[ɛ]**	**[ø]**	**[œ]**	**[ə]**	**[a]**	**[ɔ]**	**[o]**	**[u]**
F1 actual	367	396	441	620	480	582	614	705	550	393	348
F1 expect.	**282**	**311**	**391**	**548**	**405**	**519**	**544**	**765**	**571**	**411**	**327**
F2 actual	2362	1790	2152	1890	1624	1651	1425	1551	1150	993	992
F2 expect.	**2352**	**1766**	**2293**	**1928**	**1553**	**1565**	**1605**	**1479**	**1137**	**865**	**905**

Table 2 Mean formant values of German vowels produced by German native speakers

As shown in Table 2, the mid central vowel [ə] shows higher level of F1 and strikingly lower values of F2: schwa is then moved back, more open and consequently shorter and lower.

The articulatory change is at first glance surprising, because it cannot be imputed to a transfer process from the second language: indeed, schwa is not present in the Italian vowel chart.

Nevertheless, the phoneme [ə] is common in the dialectal variety used in Naples: therefore, we hypothesized the influence of the local dialect and tried to shed some lights on the possible connection between the unexpected articulation of schwa and the use of dialect. Thus, we decided to record two Neapolitan native speakers with the same map task technique used to elicit Italian dialogues[9]; results are shown in Table 3. Final unstressed vowels are often reduced to schwa in Neapolitan (Bullock, 2000); as a consequence, we found plenty of examples for our analysis.

	Neapolitan F1	**German F1**	**Neapolitan F2**	**German F2**
Schwa	**614**	614 (expected 544)	**1376**	1425 (expected 1605)

Table 3 Mean formant values of Neapolitan Italian

The results seem to indicate that the dialectal variety in use in Naples has indeed influenced the production of schwas in German: mean F1 and F2 values of [ə] deviate from expectations

[9] Information about expected F1 and F2 values of [ə] may exist, but lies outside our present knowledge.

in line with Neapolitan productions, as shown in Table 3. In fact, we would expect German speakers to produce [ə] with F1 at 544 Hz and F2 at 1605 Hz; actual realizations, on the contrary, have mean frequency values of 614 and 1376 Hz, in line with Neapolitan expectations: F1 perfectly matches Neapolitan standards, while F2 only differs for 49 Hz. Consequently, a transfer process from the L2 environment can be hypothesized in this case.

Another interesting production pattern resulting from our analysis on L2 and articulation regards the two Italian mid back vowels [o] and [ɔ] in the German part of our Interview: in this phase, interviewees spoke German, but they were asked to describe their life in Italy, the cities they had lived in, thus pronouncing many Italian words. In our examination, we noticed that the mid back vowels [o] and [ɔ] entailed in these Italian words are pronounced with formant values perfectly matching Italian means.

Language	**Word**	**Expected vowel**	**F1**	**F2**	**Resulting vowel**
German	**<Modena>**	**[ɔ]**	**567 Hz**	**1115 Hz**	**[ɔ]**
German	**<Toscana>**	**[o]**	**405 Hz**	**1093 Hz**	**[o]**
Italian	<pallone>	[o]	324 Hz	931 Hz	[u]
Italian	<allora>	[o]	283 Hz	931 Hz	[u]
Italian	<proprio>	[ɔ]	355 Hz	992 Hz	[o]/[u]

Table 4 Production of mid back vowels in German and Italian by German native speakers

However, it has to be underlined that Italian [o] and [ɔ] happen to be produced with expected F1 and F2 values only during the German-speaking phase of the recordings. The objective capability of the speaker to articulate the same Italian vowels disappears during the Italian-speaking phase (as shown in Table 1), and altered and distorted pronunciations come into being, with formant values indicating the presence of different vowels[10]. In Table 4, we provide some examples of Italian words produced during both the German and the Italian phases, underlining the different formant values and the consequently different resulting vowels.

We can say, then, that formant values of Italian mid back vowels vary in a consistent way across different linguistic contexts in our corpus. More precisely, the ability of producing formant values as expected seems to depend on the language used, thus implying that transfer processes, L2 fluency and linguistic context cannot be analyzed in a discrete way, but are of a relative nature.

3. Discussion

We can now list the suggestions resulting from this empirical analysis and discuss the possible interpretations prompted by the obtained results.

In the comparison between the German and the Italian vowel charts, some mis-articulation phenomena have come to light: firstly, formant values of Italian vowels are considerably altered, apparently due to transfer phenomena from the L1; secondly, the articulation of schwa in German appear to be influenced by the local dialect, introducing transfer

[10] Cf. expected F1 and F2 values of Italian formants in Table 1.

phenomena from the L2, too; lastly, the production of Italian mid back vowels by German native speakers is in line with expectations, but only when articulating Italian words during German productions. It seems, in fact, that German speakers are able to master the production of Italian formant values in a very precise way–but only with a restricted target: when they have to face long sentences/periods instead of few items, the same kind of vowels are not produced in the expected way anymore.

Noticeably, we recognize two different kinds of situations: in the first case, transfer phenomena take place from the L1, as for example in the production of Italian [o] and [e]: these phonemes do not exist in their short version in German: as a consequence, they are pronounced with deviating formant values in Italian.

Some other transfer phenomena, however, seem to take place from the L2: the production of schwa in German reports formant values that are typical of the Neapolitan variety of Italian. Our study, then, produces further attestation of transfer processes from the L2: to our knowledge, at the moment there are only few works on this topic (Talebi, 2013).

We hypothesize that there are two main alternatives corresponding to the above-mentioned situations:

1. With stressed vowels, as for example [e] and [o], transfer processes from the L1 take place;

2. In unstressed contexts, German speakers are used to give less importance to vowels production: transfer from the L2 language can take place.

Despite the limited nature of our results from the quantitative point of view, we observed a clear-cut distinction between full vowels – prone to transfer processes from the L1 – and the reduced vowel [ə], whose articulation in L1 seems to be influenced by L2 formant values. Stress context, then, seems to play a role in determining if transfer processes may actually take place in spoken productions: tonicity appears to be the discriminatory feature which tells L1 and L2 transfer processes apart in our results.

Another result of our analysis that should be stressed out is the importance of linguistic context at all levels: we saw that transfer phenomena can take place in some contexts and disappear in others: indeed, Italian [o] and [ɔ] are produced in the expected way in German sentences, as the focus is restricted.

Context, then, seems to be the real key to interpret transfer phenomena: whether we refer to prosodic context and refer to tonicity, or to linguistic context–with general reference to the language spoken at the moment, it seems that the native and the acquired languages are mixed in our brain in a plastic way (explicit reference is made to Cook 2003, 2004). Depending on the communicative relevance of the element we are articulating, transfer phenomena are allowed or avoided, and each time to different extents.

This interpretation further supports Natural Phonology, as a theory grounded on human needs and aims:

> Speech styles vary, and speech is used with different degrees of attention and emotion. Consequently, different degrees of difficulty–and different kinds of difficulty–are tolerated in different situations or settings.
>
> (Donegan & Stampe 1979:139)

As it can be seen, contextual variability a fundamental factor in Natural phonology. This variability is also reflected on the acquisition of foreign languages: loans, for example, are regarded as considerably complex phenomena in Donegan (1979). The degree of complexity,

however, does not result from articulation alone, but more from a merger of constraints and applicability hierarchies: processes are applied depending on the context.

> The varying applications of a natural process from language to language, from child to child, from time to time, or style to style, reveal, when compared, the implicational hierarchies along which a natural process may be limited. Although processes are universal, they do not, of course, apply identically in all situations.
>
> (Donegan & Stampe 1979:140)

Our results further corroborate this point of view: in some context, in fact, processes acquired along with the native language are hierarchically stronger than correct pronunciations in L2, and that results in transfer processes from the L1. In some other contexts, however, these processes are not that strong, due – in our opinion – to stress contextual patterns: in those case, then, acquired processes from the foreign language are preponderant and transfer phenomena from the L2 take place.

4. Conclusion

Summing up, this brief study shed some lights over the articulatory behaviour of German native speakers with Italian L2, with specific reference to vowel charts and format values.

Our results indicate that, in semi-spontaneous speech, both L1 and L2 transfer processes are admitted, but in different contexts and to a different extent. Furthermore, a complex interplay of native competence, acquired knowledge and contextual properties seems to influence the phonological system of the speakers: disentangling all these contributions should be the aim of any phonological theory dealing with spontaneous productions.

In our opinion, these phenomena can be explained fairly well in the framework of Natural Phonology, in which "naturalness is a matter of phonetic motivation, not formal simplicity" (Donegan & Stampe, 1979:141). Language, in fact, is and remains the most widespread *natural* communicative system, in the sense that it reflects human needs and abilities, despite its enormous contextual specificity.

5. Acknowledgements

I would like to express my sincere gratitude to Prof. Dr. Livio Gaeta for inspiring this paper with his work and for supporting me throughout the process. His insight and expertise greatly assisted the research; his guidance was invaluable.

I also would like to thank Prof. Dr. Francesco Cutugno for providing me with very good pieces of advice and priceless knowledge and patience.

6. References

ALBANO LEONI, FEDERICO; CUTUGNO, FRANCESCO; SAVY, RENATA. 2005. *CLIPS corpus (Corpora e Lessici di Italiano Parlato e Scritto)* [CLIPS Corpus (Corpora and Lexicon of Spoken and Written Italian)]. Online at http://www. clips. unina. it. (accessed on 03/01/2017).

ALBANO LEONI, FEDERICO & MATURI, PIETRO. 2010. *Manuale di Fonetica* [Handbook of Phonetics]. Roma: Carrocci editore.

ANDERSON, ANNE H.; BADER, MILES; BARD, ELLEN GURMAN; BOYLE, ELIZABETH; DOHERTY, GWYNETH; GARROD, SIMON; ISARD, STEPHEN; KOWTKO, JACQUELINE; MCALLISTER, JAN; MILLER, JIM; SOTILLO, CATHERINE; THOMPSON, HENRY S. & WEINERT, REGINA. 1991. "The HCRC map task corpus". In: *Language and speech*, 34 (4), 351-366.

BROWN, GILLIAN; ANDERSON, ANNE; YULE, GEORGE & SHILLCOCK, RICHARD. 1984. *Teaching talk*. Cambridge: CUP.

BULLOCK, BARBARA E. 2000. "Consonant gemination in Neapolitan". In: *Amsterdam Studies in the Theory and History of Linguistic Science*, Series 4, 45-58.

CERRATO, LOREDANA. 2006. *Sulle tecniche di elicitazione di dialoghi di parlato semi-spontaneo* [On the elicitation techniques of semi-spontaneous speech dialogues]. Online at http://www.clips.unina.it/it/documenti/2_tecniche_di_elicitazione_dialogica.pdf (accessed on 03/01/2017)

COOK, VIVIAN. 2003. *Changing the first language in the L2 user's mind. Draft of Intro to L2 effects on the L1*. Online at http://homepage.ntlworld.com/vivian.c/Writings/Papers/FXintro.htm (accessed 03/01/2017).

COOK, VIVIAN. 2004. *Bilingual cognition and language teaching. Draft of paper for talk in Taiwan*. Online at http://homepage.ntlworld.com/vivian.c/Writings/Papers/BilCog&Teaching.htm (accessed 03/01/2017).

CRAIG, C. S. 2008. *Goldwave (Version 5.23)* [Computer software]. St. John's, Newfoundland, Canada: GoldWave Inc.

DONEGAN, Patricia Jane. 1979. *On the Natural Phonology of vowels* [PhD Thesis], The Ohio State University.

DONEGAN, PATRICIA & STAMPE, DAVID. 1979. "The study of natural phonology". In: Daniel A. Dinnsen (ed.), *Current Approaches to Phonological Theory*, 126-173. Bloomington: Indiana University Press.

KIPARSKY, PAUL & MENN, LISE. 1977. "On the acquisition of phonology". In: John Macnamara (ed.), *Language learning and thought*, 47-78. New York: Academic Press.

LÖFFLER, Heinrich. 2003. *Dialektologie: eine Einführung*. Tübingen: Gunter Narr Verlag.

NAKAMURA, MASANOBU; IWANO, KOJI; FURUI, SADAOKI. 2008. "Differences between acoustic characteristics of spontaneous and read speech and their effects on speech recognition performance". In: *Computer Speech & Language*, 22 (2), 171-184.

SENDLMEIER, WALTER F. & SEEBODE, JULIA. 2006. *Formantkarten des deutschen Vokalsystems* [Formant charts of the German vowel system]. TU Berlin, Institut für Sprache und Kommunikation. Online at https://www.kw.tu-berlin.de/fileadmin/a01311100/Formantkarten_des_deutschen_Vokalsystems_01.pdf (accessed on 03/01/2017).

SJÖLANDER, Kåre & BESKOW, Jonas. 2000. "Wavesurfer-an open source speech tool". In: *Proceeding of Interspeech*, 464-467.

STAMPE, David. 1969. "The acquisition of phonetic representation. A dissertation on natural phonology" [Unpublished PhD thesis], University of Chicago.

TALEBI, SEYED HASSAN. 2013. "Cross-linguistic Transfer (from L1 to L2, L2 to L1, and L2 to L3) of Reading Strategies in a Multicompetent Mind". In: *Journal of Language Teaching and Research*, 4 (2), 432-436.

TSE, SOU-MEE. 1980. "A criticism of Stampe's Natural Phonology in light of recent findings in the study of psycholinguistics". In: *The journal of National Chengchi University*, 42, 45-52. Online at

https://nccur.lib.nccu.edu.tw/bitstream/140.119/102266/1/fb160922152302.pdf (accessed on 03/01/2017).

WIESINGER, Peter. 1971. *Phonetisch-phonologische Untersuchungen zur Vokalentwicklung in den deutschen Dialekten: Band 1: Die Langvokale im Hochdeutschen. Band 2: Die Diphthonge im Hochdeutschen*. Berlin: Walter de Gruyter.

Valentina Schettino
Department of Literary, Linguistic and Comparative studies
University of Naples "L'Orientale"
Via Duomo , 219
80138 Napoli
email: vschettino@unior.it