Netspeak: a language variety? Some remarks from an Italian sociolinguistic perspective

(Article begins on next page)

27 April 2024

UNIVERSITÀ DEGLI STUDI DI TORINO

Questa è la versione dell'autore dell'opera:

Massimo Cerruti, Cristina Onesti (2013),

"Netspeak: a language variety? Some remarks from an Italian sociolinguistic perspective".

In: Emanuele Miola (a cura di), Languages go Web: Standard and non-standard languages on the Internet, Edizioni dell'Orso, Alessandria, pp. 23-39.

Massimo CERRUTI & Cristina ONESTI[1]

(Università di Torino)

Netspeak: a language variety?
Some remarks from an Italian sociolinguistic perspective

0. Introduction

The present paper addresses a delicate issue in the Computer Mediated Communication (henceforth CMC) debate: whether Netspeak can actually be considered as a language variety. First, we will provide a brief introduction to style variation in contemporary Italian; next, languages going web are considered and style variation in Netspeak is investigated, with a special focus on the Italian context.

1. Preliminary notes about style variation

1.1. Style variation in contemporary Italian

In general terms, it is worth considering that the Italian actually spoken in contemporary Italy is a regional variety of Italian. Differently from English and from other European languages, diatopic variation (that is variation across space) has to be considered the primary dimension of variation. Broadly speaking, every regional variety of Italian has its social varieties and – as well as each of its social varieties – encompasses situational variability (Berruto 2010).

Already for some decades, contemporary Italian has undergone a restandardization process, which is caused by the mutual interrelation between spoken and written language. Such process is characterized by the acceptance in writing of peculiar traits of speech, and by the progressive acceptance of previously non-standard features into the standard variety. Markers of low varieties (belonging to very informal and sloppy speech or to the so-called 'folk' Italian)[2] are significantly emerging also in the variety of educated speakers and in a quite controlled style.

---

[2] 'Folk' Italian (Dal Negro/Guerini 2011) or 'popular' Italian (Lepschy 2002) is meant as the social variety of Italian mastered by less educated speakers.

This leads to the realization of what Berruto (1987) called *neo-standard* Italian, that is, a new standard variety of Italian. In this framework, standard Italian no longer coincides with the 'medium'-variety along the axis of formality, but it becomes often the reference point for the formal style. On the other hand, the neutral style tends to move closer to this *neo-standard* variety (see Rovere 2011). Consequently, markedness of formal styles increases and markedness of lower varieties decreases.

At the same time, regionally-marked features tend to be accepted also in written and formal contexts, giving rise to standard regional varieties (see Cerruti 2011); a tendency that seems to be rather common throughout Europe (see Auer 2011).

We can consider briefly that also Swiss Italian displays a national standard variety: it shows some differences from the Italian spoken in Italy, with some peculiar tendencies, which are related to the contact with the other three Swiss national languages and to internal dynamics (Pandolfi 2009). It is not our purpose here to enter into details of Swiss Italian; however, some Swiss studies are useful to outline further aspects related to style variation. A research on elderly people in Switzerland, for example, has showed a broader range of lexical means and of syntactic constructions in elderly people in comparison to young speakers' speech: in particular the former master a wider range of formal words and constructions than the latter (Taddei Gheiler 2005). Hence, the main differences connected to the age seem to be sociolinguistic and to correlate with the mastering of a different range of situational varieties. Crystal (2010: 231) noted as well the more conservative and formal style introduced by older people in texts of e-mails.

Interestingly enough, youth language, which is by nature very receptive to elements of spoken and informal styles,[3] shows a trend to an undifferentiated use of a single informal style (Canobbio 2005, Fusco 2007) and, at the same time, to frequent style fluctuations in some contexts, which are often related to a poor competence in managing formality (Scholz 2000, Fusco 2007).
Netspeak is recently playing a role in such processes and affecting the aforementioned Italian restandardization process.

1.2. Style variation in Netspeak

Moving to a general view on style variation in Netspeak, we found many similar claims by linguists, remarking the "robust mix of features from both informal spoken and more formal written registers" (Tagliamonte/Denis 2008: 5), "a hybrid language variety" (Ferrara et al. 1991: 10), sometimes leading to "a certain idealization […] of typical language properties, setting a benchmark

---

[3] See Pistolesi (2005), Ursini (2005), Fiorentino (2007), Gheno (2009).

against which CMC could be conceptualized as a blend or hybrid of written and spoken aspects of language" (Androutsopoulos 2011).

This was the 'first wave' of linguistic CMC studies, bringing to a sort of "Netspeak myth" (what Dürscheid 2004 has called *Mythos Netzsprache*): it denoted the popular conception of language use on the Internet as being distinct and homogeneous. As Androutsopoulos pointed out, "paradoxically, perhaps, academic work has done its part in perpetuating Internet language myths, in a manner similar to the 'unintentional reproduction' of stereotypes in language" (2006: 420). At least on the axis of formality, written language repertoires are extending to approximate the stylistic range available in spoken language (Androutsopoulos 2011).

In this regard, the position of Baron (2002: 6) is also interesting, when she speaks of linguistic *whateverism*: "The primary manifestation of this attitude is a marked indifference to the need for consistency in linguistic usages", language users being unaware of some rules. As the "whatever generation", they actually seem not to care about a whole range of language rules, leading to a "quiet revolution in social attitudes towards linguistic consistency" (2002: 2): to make but few examples, hesitation in choosing when to use *who* versus *whom* (as direct or indirect object); the case, mentioned by Baron, of Bill Gates and many other authors who could not keep straight when to use *affect* and when *effect*; high number of misspellings, missing proofreadings, etc. (see also Fiorentino; Tavosanis, this volume), which leads to the well-known CMC issues of a 'relaxed' orthography, use of emoticons and of expressive punctuation, shortened words and so on.

However, we need to avoid exaggerated assumptions about the distinctiveness of new media language. It is not possible to find a linguistic generalization that applies comfortably to digitally mediated communication as a whole, as a monolithic medium, considering in particular so different types of texts and practices present on the Internet, with a wide stylistic range in very different 'types' (in the sense of Baron 2003b) of CMC.

Moreover, linguists have to pay attention to the high speed of change: as Crystal 2010 already noted, linguistic studies in this field seem always to be out-of-date as soon as they appear. Even within a single 'type' of CMC, it is difficult to keep pace. How can we generalize e.g. about the linguistic style of emails?

Therefore, CMC studies are recently setting some recommendations and directions, reminding that we have to do with a socially-situated discourse, which cannot be considered as only influenced by the medium: Androutsopoulos (2006: 421), e.g., brings Crystal's 'variety of group practices' to the center of attention and rejects a technological determinism, shifting to the study of user-related patterns of language use.

Furthermore, "technology often enhances and reflects - rather than precipitating - linguistic and social change" (Baron 2003a: 88). Beyond some new linguistic phenomena appeared with CMC, we have to consider many pre-digital forerunners and the strengthening of tendencies, which were already attested. Netspeak has enhanced some pre-digital tendencies : abbreviations reproduce mechanisms of simplified varieties, like the language of notes; single letter respellings have predecessors ranging from African-American poetry to heavy metal record sleeves; graph-by-graph substitutions are used by some political subcultures (Sebba 2009, Shortis 2009).

2. Style variation in Italian Netspeak

The aforementioned mixing of different styles was considered a general trend in spoken Italian already before the 'web era': Berruto (1993) underlined a noteworthy disorder in the linguistic models and in mastering situational variability. Expansion of informality was firstly diagnosed for Italian by Raffaele Simone (1980), and by his hypothesis of a public sphere of action undergoing the penetration of 'private' styles as a consequence of the historical events of 1968: then the spread of a markedly private and informal style gained more domains.

As regards the specific case of style variation in Italian Netspeak, it shows some tendencies we have already considered for Italian style variation in general, regardless of CMC.

So we can summarize the main findings of the existing studies on Italian Netspeak, which are more focused on morphosyntax and textuality. They show two main tendencies: on the one hand, a tendency towards a single style, modeled on informal speech (see Pistolesi 2004: 250), on the other hand, a tendency towards the co-occurrence of markers of formal styles and markers of informal styles (see Pistolesi 2004; Moretti et al. 2004; Berruto 2005; Algozino 2011; Cerruti in press).

Just one brief example among many others[4] from a corpus of newsgroup messages:[5] *quoto questa sesquipedale fesseria. Ad imperitura memoria Joker qua di fessi ce ne sta' uno solo* 'I quote this massive blunder. To your undying memory, Joker, there's just one fool here', where the formal term *sesquipedale* ("enormous, immense") coexists with the verb *quotare*, typical of the Internet language (among other meanings of the verb, it is clearly used here as the neologism deriving from Eng. *to quote*), and with the quite informal *fesseria* ("blunder, twaddle"). The expression *Ad imperitura memoria* shows again a very high level of formality, whereas the final part of the

---

[4] See also Cicalese (2007), Bazzanella (2011).
[5] NUNC-*Newsgroups UseNet Corpora* (see below, note 8).

translation tries to convey the sloppy impact of the marked syntactic construction. Such informal effect is reinforced by superficial punctuation and orthography (see non-standard form *sta'*).[6]

With regard to a recent study about Swiss Italian (Casoni 2011), style shifting can be influenced by different functionalities related to styles. We can remark the usefulness of such shift in order to collocate what is written at a notional level of orality or conversation.

Since lexicon is one of the most sensitive fields to style variation, it is worth reporting the recent results of Algozino (2011) in order to give some empirical data about lexical variation.

After identifying a sample of sociolinguistic variables at the lexical level, Algozino has investigated whether these variables are mainly realized with formal style variants, neutral style variants or informal style variants, exploring differences between four corpora.

The corpora were consciously selected according to different peculiarities:

- LIP (*Lessico di frequenza dell'italiano parlato* "Frequency Lexicon of Spoken Italian") is a corpus of texts of spoken Italian,[7] characterized by bi-directional exchanges with free turn-taking;
- *Atheneaum* is a corpus of Italian academic writing, built up with texts produced by an Italian University in the full range of its activities, from the scientific domains to the administrative tasks.[8]
- NUNC-A and NUNC-B are two different subcorpora of NUNC (*Newsgroups UseNet Corpora*,[9] made of newsgroup messages), differing each other from the viewpoint of formality. NUNC-A includes posts on science, culture, politics, etc., while B, the more informal one, includes messages on a wide variety of topics: chitchat, pub talk, playing around etc.

Algozino has singled out 38 sets of synonyms: each synonym within a given set is associated with a different style. In other words, each synonym within a set worked as a style marker. Looking at the occurrences of such lexical markers in the four corpora, we find an expected distribution from the frequency count in LIP and Athenaeum: informal and neutral style markers prevail in spoken Italian; while formal and neutral style markers prevail in Italian academic writing, in which informal style markers are almost not attested (1%).

---

[6] *Sta'* instead of *sta* could also be the outcome of the erroneous spelling *stà* conveyed through an American keyboard.
[7] LIP is available online at: http://badip.uni-graz.at.
[8] See details at http://www.corpora.unito.it/ateneo/ateneo_.php.
[9] NUNC corpora are available online at: http://www.bmanuel.org/projects/ng-HOME.html.
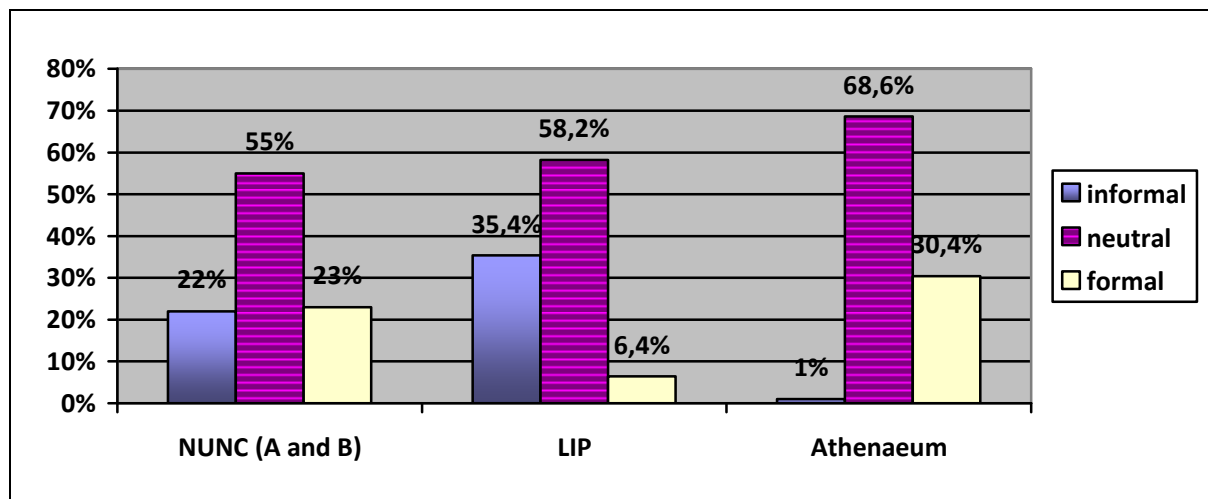
Fig. 1. Distribution of lexical style markers per corpus (Algozino 2011)

More interestingly, as for the newsgroups (NUNC-A and B), the co-occurrence of markers of formal style with markers of informal style gives rise to a similar distribution of them (22% and 23%, in the columns on the left) – differently from the clear "imbalance" displayed by both LIP and Athenaeum.

3. Netspeak: a language variety?

3.1. *In what sense do we speak of language variety?*

A crucial question that lies at the center of the debate in CMC studies is whether or not Netspeak can be considered as a language variety. To address this issue, we must first define in what sense we speak of language variety. We can speak of language variety in at least three different senses: in a narrow sense, in a broad sense, and in the sense of the so-called 'enregistered' language variety. Here we will address the issue using the term in a narrow sense; nevertheless, it's worth saying a few words about the other two senses. Let's start from the latter.

A so-called 'enregistered' language variety is conceived as "a socially recognized register" (Agha 2003: 231); in other words, a 'socially perceived' language variety. It is not necessarily identified by the co-occurrence of linguistic features, it is rather a product of language ideologies. Recently, Squires (2010) has spoken of Netspeak in these terms. She illustrates a socially shared tendency to consider Internet linguistic practices as unique and thus as constituting a distinct variety. She argues that Netspeak is perceived as a language variety on the basis of mainly two ideologies: a so-called standard language ideology, that would lead to perceive internet linguistic features as deviant from the norm; and a so-called technological determinism ideology, which holds

that technology inherently determines social and linguistic practices (see also Fiorentino, this volume).

In a broad sense, instead, a language variety is identified by differences in pronunciation and vocabulary as well as in grammar which correlates with socio-geographical or situational differences (see Trudgill 2000). It is "a member of [the] verbal repertoire" (Fishman 1972: 15) of a linguistic community. In its broad sense, the notion includes dialects, styles and registers as well as, say, sports commentary, language of television, language of newspapers, and so on. David Crystal, for instance, in his book *Language and the Internet* uses the term *variety* in this broad sense.[10] He claims that "some sort of Netspeak exists, a type of language displaying features that are unique to the Internet" (Crystal 2001: 18); a claim that has been understood by many scholars in terms of 'Netspeak *is* a language variety' (actually, as we will see afterwards, in the second edition of his book Crystal clearly claims that Netspeak *is not* a variety; see Crystal 2006: 271).

Finally, in a narrow sense a language variety is conceivable as a set of co-occurring variants. It is identified simultaneously both by a co-occurrence of linguistic features, that is variants of sociolinguistic variables, on the linguistic side; and by the co-occurrence of these variants with extralinguistic, socio-situational features, on the extralinguistic side (see Berruto 2004a; see also Berruto 2010: 229).[11] In this sense, the notion is grounded on the concept of "the linguistic variable as a structural unit" (Labov 1966; see Chambers 1995: 12-25), which is a cornerstone in Labovian Sociolinguistics.

3.2. *Language variety in a narrow sense*

Therefore, in a narrow sense, it is a set of co-occurring variants that gives rise to a certain language variety. We suggest to conceive this set as defined by the two following characteristics: 1) the set is made up of variants that statistically prevail, or even occur categorically, as compared to their counterparts; that is, given a sociolinguistic variable realized with a certain number of variants, the variant that more than others tends to co-occur with a given social feature may give rise to a language variety linked to that social feature; 2) the set is made up of variants that occur mainly in that variety, or even that are unique features of that variety; that is, a variant that co-occurs mainly or exclusively with a given social feature may give rise to a language variety linked to that social feature.

---

[10] "As Internet linguistics develops, more sophisticated models will be needed to capture all elements of the variation found. For the present book [...] I have avoided a more complex terminological system, and used the term *variety* without further qualification for all kinds of situationally influenced language" (Crystal 2001: 6).

[11] See Catford (1965: 84): "a language variety [...] is a sub-set of formal and/or substantial features which correlates with a particular type of socio-situational features".

We will soon exemplify these two characteristics. Suffice it to say, for now, that we consider these two characteristics as the real conditions to be met by a set of co-occurring variants for speaking of language variety, and that these conditions are related in a logical disjunction. In other words, we assume that we can speak of language variety if at least one of these two conditions holds true. We may arrange them as in Tab. 1.

	prevail on (/occur categorically compared to) their counterparts	occur mainly in (/are unique features of) that variety
A-type variants	+	+
B-type variants	+	–
C-type variants	–	+

Tab. 1. A language variety: types of variants

That allows to single out three different types of variants, each defined by at least one of the aforementioned conditions. A language variety in a narrow sense is therefore identified by the co-occurrence of A-type variants and/or B-type variants and/or C-type variants; in other words, variants of at least one out of these three types.

By way of example, let us consider a socio-geographic variety of Italian: the social variety of Piedmontese Italian spoken by less educated speakers, the so-called 'folk' Piedmontese Italian (see Cerruti 2011 and note 2).[12] The set of co-occurring variants that identifies 'folk' Piedmontese Italian comprises all three types of variants (we rely on data from Cerruti 2009); see Tab. 2 (sociolinguistic variables are put in round brackets).

---

[12] Piedmont is a region of Northern Italy bordering France. In Italy, common Italian speakers regularly speak a regional variety of Italian, and every regional variety of Italian has its social varieties (see Section 1.1).

A-type variants:	e.g. *genare* (imbarazzare)
	mi gena dirlo
	"I feel embarrassed about saying that"
B-type variants:	e.g. *lei* (sé)
	lei vede solo lei in concorso
	"she sees only herself [lit. her] in competition"
C-type variants:	e.g. *essere in cammino a*+INFINITIVE (stare+GERUND)
	sono in cammino a scrivere
	"I'm writing [lit. I'm on the way to write]"

Tab. 2. Types of variants, an example: 'folk' Piedmontese Italian

As for A-type variants, see the case of the sociolinguistic variable (imbarazzare), "to embarrass": its 'folk' regional variant *genare* prevails on its counterparts, e.g. *imbarazzare*, and occurs mainly in that socio-geographic variety of Italian (hence, both of the aforementioned conditions hold true). As for B-type variants, the use of non-reflexive pronouns for a reflexive meaning makes (sé) a sociolinguistic variable; its variant *lei*, for instance, prevails on its counterpart *sé* but does not occur mainly in Piedmontese Italian (one out of the two conditions holds true). As for C-type variants, the sociolinguistic variable (stare+GERUND), which is a progressive periphrasis, can be realized in 'folk' Piedmontese Italian with the variant *essere in cammino a*+INFINITIVE; a variant that does not prevail on its counterparts, e.g. *stare*+GERUND, but occurs mainly in that socio-geographic variety (again, one out of the two conditions holds true).

One last issue needs to be addressed: how many co-occurring variants are required to speak of language variety? We can maintain that a language variety "may contain just a handful of items" (Hudson 1996: 23), only provided that this is a 'handful' of *structural* items. It is worth remembering, indeed, that items of a different kind, that is variants at different levels of the language system (a phoneme, a morpheme, a lexical item etc.), *do not* have the same power to give rise to a language variety. Few structural variants may be enough to identify a language variety; whereas it may not be so for few merely lexical variants.

3.3. *A working hypothesis*

Now, what about Netspeak? Does it meet at least one of the two aforesaid conditions? Let us start by examining the first one. A question arise: if it is possible to recognize a set of variants correlated

to Netspeak, does any variant prevail on its counterparts? Actually we are not able to come up with a definitive answer, since there are very few corpus-based studies on Netspeak carried on using the concept of sociolinguistic variable. What we can do is just try to put forward a working hypothesis.

The studies grounded on the notion of sociolinguistic variable mostly deal with features often claimed to predominate in Netspeak, and call into question the fact that in Netspeak these features have higher frequencies than their standard counterparts.

Tagliamonte and Denis (2008), for example, select a set of sociolinguistic variables sensitive to style variation, and analyse the frequency of the respective variants in Instant Messaging. They investigate the frequency distribution of variants of deontic modality (*must*, *have to*, *have got to*, *got to*, and *need to*) and variants of future temporal reference (*going to*, *gonna*, *will*, *'ll*, *shall*, and the simple or periphrastic present), to name but two. Contrary to expectations, they find that variants often claimed to prevail in Netspeak *do not* have higher frequencies than their counterparts. Moreover, they find a wider range of variants than in spoken or written language, and come to describe Instant Messaging as an amalgam of linguistic features taken from different language varieties (Ferrara, Brunner and Whittemore 1991: 30 already claimed that computer-mediated dialogue "is a hybrid"):

> For every linguistic variable, IM [Instant Messaging] demonstrates a unique fusion of variants. Simultaneously, it makes use of formal variants such as *shall* and *must*; informal variants such as *will* and *have to*; and highly colloquial variants like *gonna* and *gotta* […]. It is important to point out that the type of combination of features that the quantitative analyses have revealed is not due to amalgamating different conversations from different registers. Even a single turn may contain variants of contrastive formality […]. This consistent juxtaposition of 'forms of a different feather' is the quintessential characteristic of IM discourse (Tagliamonte and Denis 2008: 25-26).

Analogous findings emerge with respect to different 'types' of CMC: among others, E-mail (see for instance Cho 2010) and Internet Relay Chat (see Paolillo 2001: 208-209, "an IRC channel […] exists at the nexus of a multiplicity of influences offering competing linguistic variants"); see also Baron (2004). Italian research has found similar results, showing a tendency to indiscriminately mix markers of formal styles with markers of informal styles (see Section 2); for instance, the quantitative analysis carried on by Algozino (2011), whose results have been briefly reported in Section 2, clearly displays this tendency. (Only in SMS it seems to prevail a tendency towards a single style, see Pistolesi 2004).

Besides, some studies show that the few linguistic features which seem instead to be more frequent than their counterparts in Netspeak are not structural features; they typically pertain

capitalization, spelling and punctuation (see Squires 2010: 482-483). Hence, in light of the above (Section 3.2), their co-occurrence may not suffice to identify a language variety.

At this point, let us take into account the second condition: if it is possible to detect a set of variants correlated to Netspeak, does any variant occur mainly or exclusively in Netspeak?

First of all it must be said that many features often claimed to be unique to the Internet not only have long been used in other written contexts (note-taking, private correspondence, telegraph messages, and so on), but also rarely occur in certain 'types' of CMC (multicapitalization or leetspeak features rarely occur in email messages, for instance); among others, see Baron (2004), Herring (2004), Dürscheid (2004), Sebba (2007), Tagliamonte and Denis (2008).

Italian research addresses similar issues. In particular, many Italian studies have shown that linguistic features which have long been used in youth language, both in oral and written contexts, are now particularly widespread in Netspeak (see e.g. Pistolesi 2005; Ursini 2005; Fiorentino 2007; Gheno 2009).

Nonetheless, the occurrence of a given feature both in Netspeak and in other contexts is one thing, and the fact that certain features occur mainly in Netspeak is another. If one reviews the existing literature on Netspeak, one may claim that features such as, say, certain lexical items, spelling variants, letters-plus-number combinations, multicapitalization, leetspeak features, abbreviations or acronyms actually seem to occur mainly or exclusively in Netspeak. (It is understood that research on whether and how many variants are used more in Netspeak than in other written or oral contexts is needed; see Squires 2010: 482-483). Undeniably, Netspeak displays somewhat typical linguistic features. But the point is that they mostly seem to pertain lexicon and graphematics (see also Tavosanis 2011, and this volume); and, again, we wonder whether – or to what extent – few features of this kind can give rise to a language variety.

Then we must take into account a frequently addressed issue, which we have not yet mentioned: with regards to Netspeak, it is questionable whether we can speak of language variety "simply because the vast diversity of users, settings and purposes […] outweigh any common linguistic features" (Androutsopoulos 2007: 280). In fact we are faced with different 'types' of CMC: E-mail, SMS, Instant Messaging, Internet Relay Chat, Newsgroups, and so on; as well as with different users: Internet language varies according to the social characteristics and relationships of its users, an undifferentiated Internet user does not exist. Crystal (2006: 271) claims that "although there are a few properties which different Internet situations seem to share, these do not in aggregate make a very strong case for a view of Netspeak as a variety". And then he wonders (*ibidem*): "if Netspeak is not a variety, what is it?".

To finally put forward a working hypothesis, we refer to the theoretical categorization of language variation according to use that we find in Berruto (1987); we depict it in Tab. 3 below:

Language variation according to use		
Language varieties (*varietà di lingua*) [clusters of co-occurring variants]		Types of language uses (*modalità d'uso*) [no clusters of co-occurring variants]
Styles (*registri*) (in the sense of Trudgill 2000: 83: "varieties […] on a continuum ranging from the very formal to the very informal")	Registers (*sottocodici*) (in the sense of Trudgill 2000: 81: "varieties that are linked […] to particular occupations or topics")	they display features of different styles and different registers; they are identified mainly on the basis of peculiar genres and types of text; cases in point: language of advertising, language of newspapers, etc.

Tab. 3. Language variation according to use, stemming from Berruto (1987)

Berruto accounts for diaphasic variation by distinguishing language varieties from what he terms *modalità d'uso*, which here we translate as *types of language uses*. The former, that in turn are divided into styles and registers, are defined by clusters of co-occurring variants, while the latter are not.[13] The here-called *types of language uses* are made up of features of different styles and different registers, and are identified mainly on the basis of genres and types of text, rather than on the basis of typical linguistic features. Berruto (1987: 159-160) takes language of advertising, language of newspapers, and, generally speaking, language of the media as cases in point.

Here we suggest to include Netspeak among these types of language uses. More precisely, we suggest to consider Netspeak as a *range* of types of language uses. We might speak, for instance, of different 'computer mediated types of language uses', and classify them on the basis of whether the communication is monologic or dialogic, and whether the dialogue is synchronous or asynchronous, one-to-one or one-to-many. We might speak, for instance, of a 'computer mediated asynchronous one-to-one dialogue', a type of language uses identified on the basis of a genre such as E-mail dialogue (which, in turn, is peculiar of a specific 'type' of CMC); or of a 'computer mediated

---

[13] In this perspective, the existence of entities which are not defined by clusters of co-occurring variants does distinguish language variation according to use from language variation according to users. More generally, on the other hand, socio-geographic varieties are strictly related to the social identity and the geographic provenance of the speaker, hence they are governed by well-established sets of co-occurrence rules more than situational varieties.

synchronous one-to-many dialogue', a type of language uses identified on the basis of a genre such as Chat dialogue; to name but two.

Notwithstanding, some of these types of language uses may partially overlap with registers: registers are characterized by linguistic features related to particular activities; and Netspeak too, as we have mentioned above, shows somewhat typical linguistic features, that occur mainly in certain 'types' of CMC.

However, as for language variation according to use, even language varieties – which are defined by clusters of co-occurring variants – are characterized by fuzzy boundaries and weak ties among variants. As is well known, establishing the existence of discrete entities is more problematic in the case of diaphasic variation than in the cases of diastratic and diatopic variation (see Ferguson 1982; see also note 12). Notwithstanding, Netspeak highlights not only that linguistic features may be very loosely tied in clusters, but also that linguistic features may happen to be not tied in clusters at all, and yet may give rise to identifiable linguistic entities.

4. Concluding remarks: item-based vs. variety-based view of language

One may wonder whether the existence of entities whose distinctive property is the amalgam of features of different varieties can undermine the cornerstone of the notion of language variety itself, which in a narrow sense is defined indeed by a set of co-occurring variants. The existence of language varieties defined by weak ties among variants serves to Hudson (1996: 45-49) as a starting point to the development of a so-called 'Item-based view of language', in explicit contrast to what he calls a 'Variety-based view of language'. In a Variety-based view, linguistic items (that is, variants of sociolinguistic variables) directly stand for a language variety; they are linked to social categories only indirectly via a language variety. In other words, it is a whole language variety that has social significance. In an Item-based view, instead, linguistic items directly stand for social categories, hence it is a linguistic item that has social significance.

From a methodological perspective, we maintain that one should follow an Item-based view of language to determine whether a set of linguistic features can identify a language variety or not. This is what we have done addressing Netspeak: we have argued that linguistic items which co-occur with a certain extra-linguistic feature can give rise to a language variety *only* if they meet certain conditions. A Variety-based view of language tends instead to take for granted the existence of a language variety.

Nevertheless, following an Item-based view of language, Hudson suggests "to avoid the notion 'variety' altogether as an analytical or theoretical concept, and to focus instead in the

individual linguistic items" (1996: 68), unless one needs "to capture generalizations that apply to very large collections of items" (1996: 48). It is a question that challenges the tricky balance between theoretical categories, which entail a certain degree of abstraction and discreteness, and empirical facts, which tend instead to be continuous.

From an epistemological point of view, we share instead the stance of those who maintain that the discrete nature of theoretical categories is not at odds with the continuous nature of empirical facts:

> The fact that reality is anything but discrete, the fact that it is fluctuating, nuanced, subject to continuous microvariability does not authorize us to think that the theoretical models that describe and possibly explain it must also be equally continuous, fuzzy, lacking strong categories, almost a one-to-one representation of reality. [It] is not a sufficient reason for abandoning the idea of working with abstract, well-defined categories. On the contrary, it should represent a stimulus and challenge to improve and refine them (Berruto 2004b: 306-307).

Stemming from Berruto (1987), we have argued for refining the categorization of language variation according to use by distinguishing language varieties from what we have termed 'types of language uses'. The notion of language variety still suits some linguistic entities, especially with respect to diastratic and diatopic variation (it is not by chance that we have addressed the notion of language variety in a narrow sense by referring to a socio-geographic variety; see Section 3.2); it is understood that in certain cases, especially with respect to diaphasic variation, co-occurrence rules are not well-established. On the other hand, identifiable linguistic entities which are not defined by clusters of co-occurring variants do exist, pertaining language variation according to use; and it is worth carving out a peculiar space for them.

References

Agha, A. (2003), *The social life of cultural value*, in *Language and Communication* 23, pp. 231-273.

Algozino, E. 2011. "Lessico e variazione di registro: un confronto tra i corpora NUNC, LIP e Athenaeum". In Cerruti, M. / Corino, E. / Onesti, C. (eds.). *Formale e informale. La variazione di registro nella comunicazione elettronica.* Roma: Carocci, 183-203.

Androutsopoulos, J. 2006. "Introduction: Sociolinguistics and computer-mediated communication". *Journal of Sociolinguistics,* 10/4: 419-438.

Androutsopoulos, J. 2007. "Style online: Doing hip-hop on the German-speaking Web". In Auer, P. (ed.). *Style and social identities: alternative approaches to linguistic heterogeneity*. Berlin-New York: de Gruyter, 279-317.

Androutsopoulos, J. 2011. "Language change and digital media: A review of conceptions and evidence". In Kristiansen, T. / Coupland, N. (eds.). *Standard languages and language standards in a changing Europe*. Oslo: Novus, 145-161.

Auer, P. 2011. "Dialect vs. standard: a typology of scenarios in Europe". In Kortmann, B. / van der Auwera, J. (eds.). *The languages and linguistics of Europe. A comprehensive guide*. Berlin-New York: de Gruyter, 485-500.

Baron, N. 2002. *"Whatever": A New Language Model?* Paper Presented at the 2002 Convention of the Modern Language Association December 27-30, New York.

Baron, N. 2003a. "Why Email Looks Like Speech: Proofreading, Pedagogy, and Public Face". In Aitchison, J. / Lewis, D. (eds.). *New Media Language*. London: Routledge, London, 102-113.

Baron, N. 2003b. "Language of the Internet". In Farghali A. (ed.). *The Stanford Handbook for Language Engineers*. Stanford: CSLI Publications, 59-127.

Baron, N. 2004. "See You Online: Gender Issues in College Student Use of Instant Messaging". In *Journal of Language and Social Psychology,* 23: 397-423.

Bazzanella, C. 2011. "Oscillazioni di informalità e formalità: scritto, parlato e rete". In Cerruti, M. / Corino, E. / Onesti, C. (eds.). *Formale e informale. La variazione di registro nella comunicazione elettronica*. Roma: Carocci, 68-83.

Berruto, G. 1987. *Sociolinguistica dell'italiano contemporaneo*. Roma: NIS (Carocci).

Berruto, G. 1993. "Varietà diamesiche, diastratiche, diafasiche". In Sobrero, A. A. (ed.). *Introduzione all'italiano contemporaneo*, vol. II, Roma: Laterza, 37-92.

Berruto, G. 2004a. "Sprachvarietät – Sprache (Gesamtsprache, historische Sprache)". In Ammon , U. / Dittmar, N. / Mattheier, K. J. / Trudgill P. (eds.). *Sociolinguistics – Soziolinguistik*, vol. I. Berlin-New York: de Gruyter, 188-195.

Berruto, G. 2004b. "The Problem of Variation". *Linguistic Review*, 21: 293-322.

Berruto, G. 2005. "Italiano parlato e comunicazione mediata dal computer". In Hölker, K. / Maaß, Ch. (eds.). *Aspetti dell'italiano parlato. Tra lingua nazionale e varietà regionali*, Münster: LIT, 109-124.

Berruto, G. 2010. "Identifying dimensions of linguistic variation in a language space". In Auer, P. / Schmidt, J. E. (eds.). *Language and Space. An International Handbook of Linguistic Variation*, vol. I. Berlin-New-York: de Gruyter, 226-241.

Canobbio, S. 2005. "Dalla "lingua dei giovani" alla "comunicazione giovanile": appunti per un aggiornamento". In Fusco, F. / Marcato, C. (eds.). *Forme della comunicazione giovanile*. Roma: Il Calamo, 33-52.

Casoni, M. 2011. *Italiano e dialetto al computer. Aspetti della comunicazione in blog e guestbook della Svizzera Italiana*. Bellinzona: Osservatorio Linguistico della Svizzera Italiana.

Catford, J. C. 1965. *A Linguistic Theory of Translation: an Essay on Applied Linguistics*, Oxford: OUP.

Cerruti, M. 2009. *Strutture dell'italiano regionale. Morfosintassi di una varietà diatopica in prospettiva sociolinguistica*, Frankfurt am Main: Lang.

Cerruti, M. 2011. "Regional Varieties of Italian in the Linguistic Repertoire". *Italian Sociolinguistics: Twenty years on* (ed. by S. Dal Negro and F. Guerini), special issue of *International Journal of the Sociology of Language,* 210: 9-28.

Cerruti, M. in press. "Variazione di registro nella comunicazione mediata dal computer: un'analisi quantitativa di dati lessicali". In Boutier, M.-G. / Hadermann, P. / Van Acker, M. (eds.). *Variation et changement en langue et en discours*, Helsinki: Société Néophilologique.

Chambers, J. K. 1995. *Sociolinguistic theory. Linguistic variation and its social significance*, Oxford: Blackwell.

Cho, T. 2010. "Linguistic features of electronic mail in the workplace: A comparison with memoranda". *Language@Internet*, 7: http://www.languageatinternet.org/articles/2010/2728.

Cicalese, A. 2007. *Da dove dgt? Chat line, testo e società*. Milano: FrancoAngeli.

Crystal, D. 2001. *Language and the Internet*. Cambridge: CUP.

Crystal, D. 2006. *Language and the Internet. 2nd edition*. Cambridge: CUP.

Crystal, D. 2010. "The changing nature of text: a linguistic perspective". In van Peursen, W. / Thoutenhoofd, E.D. / van der Weel, A. (eds.) *Text Comparison and Digital Creativity*. Leiden: Brill, 229-251.

Dal Negro, S. / Guerini, F. (eds.). 2011. *Italian sociolinguistics: Twenty years on*, special issue of the *International Journal of the Sociology of Language*, 210.

Dürscheid, C. 2004. „Netzsprache – ein neuer Mythos". In Beißwenger, M. / Hoffmann, L. / Storrer, A. (eds.). *Internetbasierte Kommunikation* (Osnabrücker Beiträge zur Sprachtheorie 68): 141-157.

Ferguson, C.A. 1982. "Simplified registers and linguistic theory". In Obler, L. K. / Menn, L. (eds.). *Exceptional Language and Linguistic*. New York: Academic Press, New York, 49-66.

Ferrara, K. / Brunner, H. / Whittemore, G. 1990. "Interactive Written Discourse as an Emergent Register". *Written Communication*, 8: 8-34.

Fiorentino, G. 2007. "La scrittura elettronica giovanile e la cooperazione. Ponti effimeri ed evanescenti o vincoli cooperativi persistenti?". In Cardinale, U. / Corno, D. (eds.). *Giovani oltre. Dal baby-boom al papy-krach*, Soveria Mannelli: Rubbettino, 259-293.

Fishman, J. A. 1972. *The sociology of language. An interdisciplinary social science approach to language in society*. Rowley: Newbury House Publishers.

Gheno, V. 2009. "I giovani e la comunicazione mediata dal computer: osservazioni linguistiche su nuove forme di alfabetizzazione". *Verbum Analecta Neolatina*, 11.1. 167-187.

Herring, S. C. 2004. "Slouching toward the ordinary: Current trends in computer-mediated communication". *New Media & Society*, 6.1: 26-36.

Hudson, R. 1996. *Sociolinguistics*, 2nd edition, Cambridge: CUP.

Labov, W. 1966. "The linguistic variable as a structural unit". *Washington Linguistics Review,* 3: 4-22.

Lepschy, G. 2002. *Mother tongues and other reflections on the Italian language*. Toronto. University of Toronto Press.

Moretti, B. / Gulacsi Mazzucchelli, E. / Taddei Gheiler, F. 2004. *L'italiano formale e la zona 'alta' della variazione linguistica*. http://www.bk.admin.ch/dokumentation/.

Pandolfi, E. M. 2009. *LIPSI-Lessico di frequenza dell'italiano parlato nella Svizzera italiana*. Bellinzona: Osservatorio Linguistico della Svizzera Italiana.

Paolillo, J. C. 2001. "Language variation on Internet Relay Chat: A social network approach". *Journal of Sociolinguistics*, 5.2: 180-213.

Pistolesi, E. 2004. *Il parlar spedito. L'italiano di chat, e-mail e sms*. Padova: Esedra.

Pistolesi, E. 2005. "Internet e il linguaggio dei giovani". In Fusco F. /Marcato C. (eds.), *Forme della comunicazione giovanile*. Roma: Il Calamo, 251-282.

Rovere, G. 2011. "Registro". In Simone, R. (eds.). *Enciclopedia dell'Italiano, vol. II*. Roma: Istituto dell'Enciclopedia Italiana Giovanni Treccani, 1228-1231.

Scholz, A. 2000. "Uso e norma della «lingua dei giovani» in base a tipi di testo informali". *Italienisch*, 44: 66-83.

Sebba, M. 2007. *Spelling and Society. The Culture and Politics of orthography around the world*. Cambridge: CUP.

Sebba, M. 2009. "Sociolinguistic approaches to writing systems research". *Writing Systems Research*, 1/1: 35-49.

Shortis, T. 2009. "Revoicing Txt: spelling, vernacular orthography and 'unregimented writing'". In Wheeler S. (ed.). *Connected Minds, Emerging Cultures: Cybercultures in Online Learning*, Charlotte: IAP, 225-246.

Simone, R. 1980. "Parlare di sé". In Galli della Loggia E. et al. (eds.). *Il trionfo del privato*. Roma-Bari: Laterza, 191-230.

Squires, L. 2010. "Enregistering Internet Language". *Language in Society*, 39/4: 457-492.

Taddei Gheiler, F. 2005. *La lingua degli anziani. Stereotipi sociali e competenze linguistiche in un gruppo di anziani ticinesi,* Bellinzona: Osservatorio Linguistico della Svizzera Italiana.

Tagliamonte, S. A. / Denis, D. 2008. "Linguistic ruin? LOL! Instant messaging and teen language". *American Speech*, 83/1: 3-34.

Tavosanis, M. 2011. *L'italiano del web*. Roma: Carocci.

Trudgill. P. 2000. *Sociolinguistics*, 4th edition. Penguin: London.

Ursini, F. 2005. "La lingua dei giovani e i nuovi media: gli SMS". In Fusco F. / Marcato C. (eds.). *Forme della comunicazione giovanile*. Roma: Il Calamo, 323-336.