International Journal for Dialogical Science
Fall, 2007. Vol. 2, No. 1, 133-161

ON THE NOTION OF VOICE: AN EXPLORATION FROM A PSYCHOLINGUISTIC PERSPECTIVE WITH DEVELOPMENTAL IMPLICATIONS

Marie-Cécile Bertau
Universität München, Germany

ABSTRACT. The notion of voice is explored in two steps. Firstly, Bakhtin's understanding of voice is sketched. Secondly, a psycholinguistic perspective is developed where voice is first of all a concrete auditive-vocal event. Five key concepts are used to describe the phenomenon: indexicality, body, intonation, imitation, and internalization. Indexicality refers to voice as index means for speaker/listeners, pointing to the actual shared situation of communicating persons; further, voice indexes the speaking body which is itself socio-culturally shaped. Intonation is viewed as being deeply socially and dialogically shaped. Imitation and internalization are explored in regard to voice acquisition in ontogenesis. Mutual imitation of child and caregiver in early communication and speech acquisition form an incessant movement from the one to the other, intermingling the voices of both persons. Eventually, the voice of a specific other leads the internalization process, voice being its "semiotic mechanism". As a result, voice is thought to be a meaningful, perceivable and experienced form tied to another person. This form serves as a powerful mechanism of internalization: vivid materiality becomes a psychological process.

Keywords: voice, intonation, imitation, internalization, dialogical psycholinguistics

The notion of voice is fundamental to the theory of the dialogical self (Hermans and Kempen, 1993; Josephs, 2002; Hermans & Dimaggio, 2004). A brief look at the notion's usage in this theory hints at some core ideas related to it, the most important one being voice as embodied entity.

In its connection to therapeutic work, the notion of voice is closely related to processes of change, to the development of new and different positions in the self. The spatialization of self (Hermans, 1996) allows for simultaneously different positions, and for movement between these positions. The I moves in this space, having the capacity to "endow each position with a voice" (Hermans, 1996), thus establishing dialogical relationships between positions. Hence, in "voice" it is the process of *giving* a voice, and through this, to come into a process of change, that matters; voice and position are the basic notions constructing the space of Self, its perspectivity, its stories, its

AUTHOR NOTE.. Please address correspondence about this article to Marie-Cécile Bertau, Institut für Psycholinguistik, Universität München, Oettingenstraße 67, D-80538 München. Email: bertau@psycholinguistik.uni-muenchen.

coherence (see e.g., Raggatt, 2006). Movements in the self are conceived either as centrifugal (multiplicity of positions, discontinuity and innovation, risk of fragmentation), or as centripetal, with emergent meta-positions (continuity and stability, risk of rigidity). These movements are in constant tension and complement each other (Hermans & Hermans-Jansen, 1995). The processes of voicing or silencing can be seen as carrying these movements. Therefore, in the work of Hermans, I would underline the *generating* character as a basic feature in the concept of voice.

Josephs (2002) illustrates quite clearly the constructional process of a voice in the dialogical self in observing the formation of a voice she calls "the Hopi in Me". Located neither in the person nor in the culture but as an emergent process between person and culture, voice develops in their contact zones. A voice is then " not a role we play" on the background of a "true self", it is not a fixed trait of the personality. Coming to a positive definition, Josephs (2002) writes:

> A voice is rather an emotionally grounded and personally constructed – in short: a meaningful – focus on one's life in the here-and-now. Thus [...] the range and characteristics of the voices populating the self are in priciple unlimited, and also unpredictable by anybody but the person her- or himself. (p.162)

In dialogical self theory (DST) the generating character in the notion of voice is taken as a fruitful theoretical heuristic in understanding diverse psychological processes and entities as "voice". This path is taken explicitly in Dimaggio and Hermans (2006, submitted) when they describe the dialogical self as an entity made up of a multiplicity of parts named either "voices" or "positions" or "characters"; each of them function as a partly independent agent generating specific memories, thoughts, and stories. Thus, following Dimaggio and Hermans, "voice" refers to an *agentic* starting point for a message addressed to any person, or to another part of the self.

Hence, it is not only the possibility of movement and simultaneous multiplicity that is interesting in the notion of voice, but also the aspect of independence and agency. It is with these very notions that Stiles (1999; et al, 2004) is able to conceive experience as embodied rather than as a mere cognitive representation. Departing from the heuristic metaphor of voice, Stiles (1999; Stiles, Osatuke, Glick, & Mackay, 2004) arrives at a literal understanding of voice, where the internal multiplicity can be externally heard and empirically analysed. In these researches, voices are firstly seen as internal, developing as traces from lived experiences, incorporating expressive, experiential, and interpersonal elements (Osatuke et al, 2005), *then manifesting* outside. Starting with the hypothesis that each of these manifesting voices sound different from each other, the authors demonstrate convincingly the identification of different voices within a person, characterizable by distinctive names and by a qualitative description of

vocal and personality features (Osatuke et al., 2004; et al, 2005).[1] Consequently, "voice" comes to be a clearly embodied entity, underscoring "the physicality of psychological self" (Osatuke et al, 2004, p.252). However, these investigations are done in the perspective of an already formed dialogical self, where new voices emerge, or internal ones manifest themselves in certain ways. From a developmental point of view the question remains how voices are formed in ontogenesis, or how the capacity of voicing is acquired.

The following considerations address this ontogenetical issue, and propose an exploration of some of the notion's facets, including a historical approach in order to clarify implicitly transmitted meanings. This is approached relying heavily on the Russian-Soviet idea of voice – not only because Bakhtin is a privileged point of departure for many scholars in the dialogical sciences, but also because this idea shows itself as rooted in a certain conception of *language.* The core arguments in this paper are made from a psycholinguistic point of view, explicitly focused on language as an event taking place between and inside people. Thus, linguistic as well as psychological dimensions of "voice" are addressed.[2]

Bakhtin's Notion Of Voice – A Sketch

As mentioned, many current writers in the dialogical sciences refer to the works of Mikhail M. Bakhtin. In order to situate the concept of voice in its socio-historical context, it should be noted that Bakhtin himself was part of a current of ideas emerging in Russia at the turn of the 20th century and further developed in the first decades of the new century in the socio-political context of the Soviet Union. Here, it is not only Bakhtin's relationship to Voloshinov and his way of conceptualizing language and consciousness that is to be mentioned (Voloshinov, 1973), but also some of the main precursors of dialogical thinking: Dal', Potebnya, and Yakubinsky (see Ivanova, 2003; Jakubinskij, 2004; Meng, 2004; Naumova, 2004; Romashko, 2000).

The specificity of the Russian socio-cultural and linguistic context at the end of the 19th century is diglossal: a deep dichotomy existed between the language of the church (Church Slavonic) and the language of everyday communication (Russian). This situation polarized written and oral language, the world of authority and power and the world of the people, and this was taken as the opposition of dead (stiffened) and alive or vivid (see Romashko, 2000). The notions of vividness, dynamism, movement – also in reference to Humboldt's *energeia* (e.g. Potebnya in Naumova, 2004) – are important to the understanding of the Russian ideas on language and thinking of that time. In Vygotsky (1987) this notion is still present and it runs like a thread through his work. It

---

1 A similar idea underlies the analysis of thinking-aloud-protocolls of problem solving persons in Bertau (1999), whereby intrapersonally distinct speaking roles could be identified.

2 For a linguistically focused approach addressing especially the issue of form, see Bertau (2007, in press).

is also, of course, found in Bakhtin's notion of dialogicality which has two antipodes: the official language characterized by fixed meaning without any ambiguity, and the poetical language which is in Bakhtin's view oriented towards the norms of an official culture. In contrast to these static, monological languages, it is the language of prose that realizes and permits ambiguity, movement, and different positions from where the "sole truth" of the official world is questioned. Here, two positions of meaning, two accents of values, two kinds of consciousness meet: there is bivocality (see Lachman, 1982).

As far as I can see, in Bakhtin "voice" is used as a vanishing point with its start in the notions of word, utterance and answer. From these notions, Bakhtin arrives at voice, thereby describing the foundations of language as a dynamic structure of acts of answering: every utterance is an answer to preceding utterances, every act of comprehension is related to an attitude towards answering, and every utterance is produced in anticipation of an answer (Bakhtin, 1986, pp. 69, 91). For Bakhtin, a living language is not conceivable outside these relations with acts of answering, outside the dialogical movement.

The same attitude relates consciousness with voice. In Dostoyevsky's characters, Bakhtin observes consciousness in continuous dialogues of voices, internal as well as external, representing "the whole person" and having a special "density" and "resilience" (Bakhtin, 1984). Voice becomes the indicator of the "essential characteristics of consciousness":

> ...for Bakhtin, dialogue is an expression of the essential characteristics of consciousness, which unite it with external, also dialogical existence [...] it is the concrete psychological embodiment and measure of the social quality of consciousness. (Radzikhovskii, 1986-87, p. 18)

Thus, Bakhtin does not view consciousness in itself (as viewed by traditional psychology) but seeks existential characteristics for it which he finds in the dialogue of voices. The social aspect of consciousness "consists in two minds addressing one another internally" (Radzikhovskii, 1986-87, p. 21). Because all reality is interpersonal communication between voices, consciousness is a voiced internal dialogue. It is from the perspective of the spoken word that Bakhtin arrives at a conception of voice which has to be thought of in terms of twofoldness: the "bivocal word" (Bakhtin, 1984) is the point of departure, not singleness. Interestingly enough, there is not only no voiceless word but the quality of the contained voices oscillates between near and far, impersonal and familiar:

> Each word contains voices that are sometimes infinitely distant, unnamed, almost impersonal (voices of lexical shadings, of styles, and so forth), almost

> undetectable, and voices resounding nearby and simultaneously. (Bakhtin, 1986, p. 124)

In these different distances one can recognize the realm of typified and typifying language uses to which all speaking is compelled because “we speak only in definite speech genres” (Bakhtin, 1986, p. 78), however “flexible, plastic, and free” they may be. Otherwise speech communication and the exchange of thoughts would be “almost impossible” (Bakhtin, 1986, p.79). I would like to hint at the “resounding” quality of the nearby voices, whereas the far ones are “almost undetectable” - a theme to be picked up further on.

A voice seems to have the function of a carrier: voice carries the speaking subject out of himself, decentering and orienting him toward the other(s) (both face to face, and general social others), supporting and leading the contact. What a voice carries and expresses at the same time is that the utterance is as well “mine” as “other's” (Bakhtin, 1986, p.89). The speaking subject matters only as a decentered and therefore “twofold” subject, endowed with a voice which carries at least two “tones and echoes” belonging to the uttered word. Voice carries the individual expression of contact with the other which is always mingled with some alien components. It supports the necessary multiplicity belonging to the living language in the word. Voice carries consciousness, manifests it. One voice, one consciousness “says” nothing, what is needed are at least “two”:

> A single voice ends nothing and resolves nothing. Two voices is the minimum for life, the minimum for existence. (Bakhtin, 1984, p. 252)

Bakhtin's point is about the “individual manyfoldness of voices” (Bachtin, 1979, p.157-159). And in this multiplicity there is movement and life, both ideas serving the counterpoint to monologism. The individual manyfoldness of voices is grounded in the social manyfoldness of speech (Bachtin, 1979, p.157), so one can assume that an individual voice is always manifold. It has a multiplicity of expressions, corresponding to the social language needed in the actual situation. This amounts to saying that, for Bakhtin, voice is not a completely individual phenomenon, on the contrary, it always transmits the typifying character of spoken national and social language, and of genres.

Therefore, the concept of voice means the construction of a social person. The ways an individual speaker may express his social identity (profession, gender, social status, age group, ethnicity etc.) include virtually every linguistic contrast: e.g. lexical and syntactical choices, but also on a paraverbal level, intonation, physical voice quality and variations in fluency (see Keane, 2001). So, the physical voice quality is one of the features of the social identity. And, because there is always more than one voice in any word, the concept of voice also means the agonistic interaction of voices. In Bakhtin's notion of heteroglossia differences are what matters. Here, voices are juxtaposed against

one another, even within the discourse of a single speaker. This kind of struggling interaction between voices militates against any notion of harmony one could naively relate to dialogicity – a point underlined by Hermans and Kempen (1993), too. For Bakhtin, difference is the necessary prerequisite of movement and dialogue, the prerequisite of life.

Depicting the cultural-political context of the 1920's in Russia, Gasparov (1982) states that Bakhtin's fundamental viewpoint was the "pathos of the expropriation of the other's word". The "new artist" of the "new (revolutionary) times" discovers that his linguistic material is completely well-thumbed. Every word was used innumerably by others. From this point of departure, the artist must find a way to express his own thoughts in an alien, inherited language. Linking this Bakhtinian experience with the situation of Russian diglossia and with the new possibilities of the first years after the Revolution, one may claim that Bakhtin and his main dialogue partners, Voloshinov, Medvedyev and Pumpyansky (see Meng, 2004), themselves faced the necessity of finding and forming their voices out of the existing traditional ways of speaking and writing. Due to their transitional situation they were quite sensitive to these ways *as ways,* to the fact that they are only one of many possibilities of expressing reality, belonging to specific ideologies. Bakhtin's dialogical thinking is itself a struggle for formulating an own standpoint which will not be as dogmatic about truth and monological as the inherited one, but includes manyfoldness, i.e. multiple voices. Thus, the socio-historical situation where "voice" emerges is related to a certain pathos of the praise of liveliness, and of the liberation of the individual, allowing for his and her plurality.

A Psycholinguist's Perspective On Voice

Generally speaking, a psycholinguistic point of view is determined first of all by the aspect of processuality of language and language related phenomena such as thinking, memorizing, meaning and understanding. This processuality is seen in actual situations and contexts as microgenesis and as actual genesis, in development as ontogenesis and follows Vygotsky's genetic principle which holds that to understand a phenomenon it is necessary to study its development (1987, chap. 1, 5). With the relationship between speech and thinking being at the core of psycholinguistic questioning, it is almost natural to arrive at a dialogical view, for this relationship addresses the links between a self and the world, understood as other selves, and their common actions. Nevertheless, it is perfectly possible to study language processes in a monological way, to assume a self-contained cognition uttering ready-made language structures for another, to be processed in a similarly self-contained way.[3] Applying the theory of the dialogical self to psycholinguistic issues does not only mean to open these

[3] See O'Connell & Kowal (2003) for an extensive criticism in that vein.

closed minds, to see co-constructing processes in speaking and thinking, it also allows and necessitates a radical change in the view of language itself. Addressivity becomes the foundation of this language conception, giving its specific relational form to all thought and speech processes (see Bertau, 2004b), foremost in development. It is this issue which is taken up in Bertau (2004a) with the model of phonicity, where the audible voice of the mother is the starting point for the infant's dialogical self-development, as I discuss later on in this article. So, voice is part of the fundamental concept of addressivity, it is the privileged way humans express their addressivity – which is not to say that the voice already utters words, but that it is first of all a concrete audible event between people in the external world.

From this point of departure two main issues should be addressed in the following: firstly, a description of the notion of voice, integrating different approaches and organized around the concepts of indexicality, body, and intonation. These concepts belong to a voice as a product of a development. The first three concepts are supplemented by two further ones: imitation and internalization accounting for the ontogenetically developed voice. These concepts point to the second issue, corresponding to the question of development which could be termed as follows: How does an internal I-position develop out of external experiences with audible voices? The aim is to arrive, via the five concepts, at a workable construct which is to be elaborated by further empirical work, especially in the field of acquisition. The general hypothesis is that acquiring language and developing a self are deeply related processes, one building link of this relationship being the voice.

Features of voice

Indexicality. The term "indexicality" refers to the context-dependency of natural language utterances, which can include various phenomena including, e.g., regional accent, indicators of verbal etiquette, referential use of pronouns, demonstratives, deictics, tense. The verbal ones are investigated by linguists, implicitly giving a paradigm of indexicality in these forms, mainly pronouns and deictics. But, as Laver (1975) writes:

> ...just the fact of speaking and allowing the other participant to hear the sound of one's voice, regardless of the actual content of the utterance, provides the listener with some information he needs to reach some initial conclusions about the psychosocial structuring of the interaction. (p. 221)

What Laver (1975) terms "phonetic behavior" is important to the participants when they construct a working consensus for the beginning interaction. The features of the voice serve as an orientation.

> When a person speaks, he reveals often very detailed indexical information about his personal characteristics of regional origin, social status, personality, age, sex, state of health, mood, and a good deal more. [...] As listeners, we infer these information from phonetic features such as voice quality, voice-dynamic features such as control of pitch, loudness and tempo, and from accent, as well as to some extent from features of linguistic choices made by the speaker. (Laver 1975, p.221).[4]

A first characteristic of indexical expressions is their co-presence with what they stand for. So, contiguity, not similarity or conventionality is defining for indexicals. Related to this is the fact that they give little or no description of their referents, they function as link to their context, not as designators of objects and properties. Therefore, as a second characteristic, indexicals are closely associated to gestures, such as pointing and showing or handing over (Hanks, 2001). Co-presence and gesture relation stress the fact that indexicals are anchored in a bodily dimension of language. A third characteristic of indexicals is that "they systematically shift in reported speech" (Hanks, 2001, p.120). This point is interesting insofar as the phenomenon of reported speech is closely related to the notion of voice: it is the voice that shifts insofar as it changes perspective and authoring, which is manifested at the phonetic level, too (Tannen, 1989). Finally, the function of indexicals is to direct the addressee "to look, to listen, to take an object in hand" (Hanks, 2001, p.119). Precisely this embodied directing of the other is found in Karl Bühler's (1934/1990) theory of indexicals.

Bühler (1934/1990) associates indexicals as deictic words with the so called deictic field related to perception, in contrast to the symbolic field associated with naming words. Relating both fields, Bühler states that the demonstrating word *individuates* what is named ("that tree"). Bühler differentiates four forms of deixis in the deictic field from which the I-here-deixis will be picked out, for its hints to voice. Departing from the questions "Who is there?" posed behind a closed door and "Where are you?" posed in the dark, Bühler analyses the answers "I" and "here". Bühler terms "I" an "individual signal", and "here" a "positional signal" (1934/1990, p.110). Identifying "the place or the person involved" is done on the basis of the *sound*. This sound reveals the individual character as well as the origin of what is expressed. So, for Bühler the core function of the primal "here-word" is to direct the gaze to the position of the speaker. The primal "I-word" does more: it not only demands to seek the speaker with the eyes but also urges the listener to aim at the speaker with what Bühler calls "a physiognomic gaze."

---

[4] Laver is right to put the "linguistic choices" at the end of the list. A. Mehrabian showed as early as 1972 that only 7% of words have influence on communication, in contrast to 38% for voice and 55% for nonverbal, bodily communication.

These ideas can be followed up in Bühler's analysis of the pronouns "I" and "thou" as indexing persons in a speech drama, therefore not designating anything. And in this, they individuate speakers and assign them a position.

> The phonologically imprinted, formed structure ich (I) [...]. resounds with the same phonological form from millions of mouths. It is only the vocal material, the auditory shape that individuates it, and that is the meaning of the answer I given by my visitor at the door: the phonematic impress, the linguistic formal factor in his I points out the vocal character to me, the questioner. (Bühler, 1934/1990, p.129)

Following this analysis, Bühler deduces as highly interesting that

> the form of something is there to the end of pointing out an idiosyncrasy in which the form is realized. (Bühler, 1934/1990, p.129)[5]

The phonological form has to point to the specificity of the material: which is sound, individually uttered sound from a certain position. So indexicality of voice means to Bühler a *turning toward* the heard speaker with a "physiognomic gaze": recognizing him/her as him/her in this specific time and place, at this certain position. Voice directs the other to an individual which is to cognize and recognize. The very possibility of understanding uttered words is in Bühler related to the positioning of the person.

With the concept of indexicality one is immediately thrown into a context of time and space surrounding speakers/listeners moving toward one another and toward the indexed actions and objects in that time and space. Actions and objects are shown, are brought into "horizons of relevance" (Schütz, 1982), that is, certain objects and actions are constructed as relevant in respect to certain of their aspects for an actual situation involving concrete persons. But first of all it is the person who shows herself the other *as* a certain one: in this, the indexical process of voicing is discovering. "Indexical claims" (Laver, 1975) shape and constrain the detailed relationship of speakers/listeners. That is, we first have to show and see each other before we can exchange any verbal content. The suggestion made here is that the uttered voice is an important index. Leaving aside social languages and genres, it first of all points to the immediate context, time and space and actual participants.

Regarding interactions that are not face to face, I would claim that they need a basis, the live experience of face to face encounters which are then transposed,

---

[5] For this sentence is central to the understanding of the functioning of voice as form it should be given in the original: "die Form eines Etwas ist dazu da, auf die Besonderheit des Stoffes, an dem die Form realisiert wird, hinzuweisen." (Bühler, 1934/1982, p.113)

transformed and abstracted. Skilful use and comprehension of indexicals outside the bodily co-presence in time and space of the interactants (on the telephone, in letters, in e-mails), are all the more elaborated as they can rely on a solid foundation of interactional practices which include the ability to take the perspective of the listener. Detachment of shared time and space requires precisely the imagination of the other and of his/her communicative-cognitive possibilities. This underscores that any abstracted, detached competence remains rooted in the experienced situation, anchored in the experience of interacting persons, gaining its very comprehensibility from there.

Body. Stressing the anchoring of any language practice, however abstract it may be, in a common, lived and shared experience, which is at least legitimized by the developmental point of view on phenomenon taken as fundamental, I arrive at "voice" from a bodily experience. Voice is a central notion to dialogical psycholinguistics for the very reason it connects speech with body and emanates from this body. The uttered voice shows, indexes the uttering body – as an individual (gender, social status, age etc.) and as a position (sitting there, coming in front of) – and leaves it as a medium of generalized, inter-individual signs, not belonging to any person (see Bakhtin, 1986; Voloshinov, 1973). Voice refers to a physical event that is never mere physics but always includes assigned meaning: meaning as related to verbal signs, and meaning as related to all embodied expressions of humans, themselves socio-culturally determined. So, "voice" is a vocal-auditory event, and it is a concept belonging to a certain socioculturally constructed way of expression. The uttered voice is absolutely individual, coming from a unique body, but this body is located in specific sociocultural contexts and has a history of actions, movements, labels, etc. So, the voice, too. As for every human expression, the voice is individual and societal, both aspects being the facets of a wholeness, and staying in contrast to "natural" (Voloshinov, 1973, p.34).

This understanding expands to non-spoken, written language where the ways of expression are detached from their vocal-auditory dimension, but are nevertheless not completely disembodied. The history of alphabetic writing and reading shows the transformation from an embodied to a disembodied practice, as well as the shift from vocally performed written language to its silent production and perception (Saenger, 1989; Coulet, 1996). This shift into disembodiment suggests a path to the assumption of the monologism of written language (Bakhtin as paradigm). However, from the perspective of the general hypothesis set here, concerning the acquisition of language and the development of a self, written language has an equal status as spoken language. In both, a plurality of voices has to be acquired, belonging to specific genres and languages as well as to the individually developed perspectives on world. Thus, written language is also to be thought of as dialogic and manifesting different voices.[6]

---

[6] This line of idea is taken up in an ongoing project investigating how students acquire the

In saying that the voice is a vocal-auditory event we refer to the double-sidedness of voice perception which may be one of the reasons why humans privilege voice as the medium of verbal communication: voice is perceivable both by the speaker himself, as proprioception, and by the listener - in that respect it is different from gaze; it is a concrete sensitive event, a means to touch the other over space, and as such it encourages transposing and abstractness. This is my proposal for what happens in ontogenesis: the child moves from the voice of its mother as a bodily experience (analogous to her touching and handling) to her voice as medium of signs. Meaning is always there, and always socioculturally shaped, first of all addressed. So, voice offers a meaningful structure in so far as it is directed toward somebody. Body and voice are inseparable. Voice refers to the body it comes from, and the kind of body shapes the quality of voice. Both are social and individual phenomena, manifesting the relationship and tensions between these two interdependent sides.

Marcel Mauss (1936/1999) was the first one to acknowledge clearly that nothing in our bodily expressions is natural, but is rather, specific to cultures and societies and even specific to generations in societies. With the term *techniques du corps* (techniques of the body) Mauss refers to the ways humans use their body, how they hold themselves, how they move, lay down, sit, stand, go, swim etc., even breathe. In his enumeration of the body techniques Mauss lists the techniques of giving birth, where, for instance differences in handling the newborn are worth noting. This is followed by the techniques of childhood, where questions of how and when to hold the baby, ways of breast feeding him/her or not, ways of stopping breast feeding, ways of laying down babies – just to mention a few – matter for the formation of the body morphology and of the attitude of the person toward himself. All forms of touching and handling are saturated by sociocultural meaning and are a means of transmitting these meanings.

Voice plays an important role in raising children. Despite not being in my opinion a technique itself in Mauss' sense, it is a necessary part of the techniques used with babies and children in that it accompanies, structures and rhythmifies all the handlings and touching of nursing. The voice stresses a certain quality of the caregiver's action: slow, smooth, rapid, impatient etc. And it is not surprising that almost all cultures have developed a so called baby talk where, besides semantic and syntactical features that reduce complexity, it is the voice quality of the caregiver that matters: it is at a higher fundamental frequency, slower and with clear intonation; moreover, the caregiver uses a lot of repetitions[7] – a Bakhtinian social language determined by the age

---

academic voice needed to write an abstract by analyzing the stages in the writing-correction-revision-process (Karsten, in preparation).

[7] See Snow (1977) for the first description of baby talk. Repetitions used in infants are to be seen as patterned structuring of time, giving as such orientation and security; they are semantically *not* redundant.

of the listener, with different genres, for example how to speak with the infant when breast-feeding, or when commenting on his/her actions.

Body means orientation in space, wherein "space" is to be understood as socioculturally constructed and organized. In turn, orientation in space means position, and this is perspective: first of all an attitude toward the other and the world, developing from the *techniques du corps*, i.e. from the socioculturally meaningful ways one is held toward other and the world. An emotional-cognitive perspective is acquired together with a body position, from where things are seen in a certain way and from which one can tell certain stories and feel certain feelings. So, the position and its perspective uttered in a voice are closely related to early body experience shaped and formed by others. If one assumes that any perspective and its position uttered in a voice develop out of the relationship others express towards self, one must include a pervasive affective attitude. The expression of such an attitude will then be part of thinking and speech; Josephs' (2002) claim for an emotional ground in voice meets this reasoning. Consequently, cognitive processes are inseparable from affective attitudes which can be expressed in a voice. Vygotsky stresses the need to see thinking as not being isolated from affect:

> Every idea contains some remnant of the individual's affective relationship to that aspect of reality which it represents. (Vygotsky, 1987, p. 50)

It only needs to be added that the individual's affective relationship is given and shaped by other's voices as expressions of their socio-individual perspectives.

Intonation. What has been said so far on body and voice is even more clearly to be seen in the nature of intonation. Because of Voloshinov's explicit and radical standpoint on the sociality and addressivity of any utterance and of any word from which a *social essence of intonation* is concluded, the concept of intonation developed hereafter will be based on this author.

Voloshinov, one of the members of the so-called Bakhtin Circle, grounds his notion of language on the idea that the utterance is generated by an experienced extra-linguistical situation (Voloshinov, 1981a, pp. 188-191). So, Voloshinov arrives at language from the outside, so to speak, and he will stress this approach, maintaining the links of verbal and extra-verbal parts. In that, the word does not mirror the extra-verbal situation, nor is this situation to be thought of as an external cause of the utterance, but the word *accomplishes* the situation, makes an evaluation of it (Voloshinov, 1981a, pp. 190ff.). To the relation of utterance to situation Voloshinov adds the relation of utterance to listener. Any utterance is conceived in regard to a listener, i.e., to his/her comprehension and answer, as well as in regard to an evaluating perspective of this listener. The utterance is therefore always directed to the other, to the listener, and this

leads Voloshinov to take social and hierarchical relations between the interlocutors into consideration (Voloshinov, 1981b, p. 298).

The notion of intonation is developed in the investigation of the form of the utterance. Voloshinov first states that an utterance without words would still have "the sound of voice" which is intonation (Voloshinov, 1981b, p.304). If even this were absent, the gesture would remain.[8] These last kinds of instances manifest the materiality of communication, a subject extensively developed in Voloshinov (1973).

> *En dehors de l'expression matérielle, il n'existe pas d'énoncé, il n'existe pas davantage d'affect*. [*Outside of material expresssion utterance does not exist; nor does affect exist*.] (Voloshinov, 1981b, p.304; emphasis in the original)

Voloshinov distinguishes three fundamental elements which organize the form of the utterance: first, intonation, which is described as the "expressive timbre of a word"; second, choice of word; third, disposition of the word within the utterance. Intonation is emphasized in that it "first of all" relates the utterance to the situation and to the audience. Besides, intonation plays the first role in the construction of the utterance, i.e., the second and the third aspect of form are built *as a consequence* of intonation (Voloshinov, 1981b, p.305). Intonation itself is determined by the situation and the audience. Voloshinov illustrates this aspect with a passage from Gogol where it is shown that a sudden, violent change in intonation occurs at the moment that the situation and the audience of the utterance are modified. Intonation is therefore the phonic expression of the social evaluation ("l'expression phonique de l'évaluation sociale", 1981b, p.305). Thus, as speakers/listeners we take an evaluative attitude toward the situation and toward one another, giving value accents which are ideologically shaped. Communicating is first the expression of a certain attitude which gives all utterances a certain accent. And attitudes form intonation, which is first an evaluation of the situation and of the audience ("une évaluation de la situation et de l'auditoire", Voloshinov, 1981b, p.307), in turn calling for the adequate word, and assigning a certain position to this word in the utterance.

The function of intonation of voice is seen by Voloshinov as similiar to the carrying function as developed in my reading of Bakhtin. The features of flexibility and sensitivity facilitate its use and make it pervasive. This supports the claim for the centrality of the notion of voice for a dialogical psycholinguistics.

---

[8] This, now from another side angle, corresponds again to Mehrabian's (1972) dictinctions among words, voice, and nonverbal communication.

> L'intonation joue le rôle d'un conducteur particulièrement souple et sensible au sein des rapports sociaux... [*intonation plays the role of a particularly flexible and sensitive leader within social relationships*] (Voloshinov, 1981b, p.305)

The social aspect of intonation is underlined in the way Voloshinov depicts its comprehension. It is only understood when one is familiar with the implicit evaluations of the social group in question, be it a family, a social class, a nation, an epoch. And in the process of producing intonations, Voloshinov draws on the addressivity of any utterance. In the case the speaker can suppose a "chorus of support" in his audience, his/her intonation will be vivid, creative, rich in nuances and self-confident. On the contrary, in the case of a lack of support, "the voice will brake", its richness in intonation will be reduced (Voloshinov, 1981a). So, what results here clearly is the deeply social character of intonation, more precisely its dialogical character, being in its features dependent on the other as addressee (who can be actually or virtually present, as Voloshinov notes).

Acquisition of Voices in Ontogenesis. In Bertau (2004a, c) I claimed that the voice plays a significant role for the development of the dialogical self. The voice of the caregiver supports and leads the development of the infant and child from diffuse social acts to clear mutual exchanges. In regard to the importance of this primary, real, embodied auditory-vocal event coming from a certain person and addressed to another certain person, the development of voice in ontogenesis was grasped in the model of phonicity. Based on early dialogical structures beyond verbal language between mother and infant (Bruner, 1983), and supplied with important steps for developing intersubjectivity (Rochat et al, 1999; Akhtar & Tomasello, 1998), the child enters speech acquisition as an already dialogical being, aimed at mutual and addressed exchanges (Lyra, this issue). That is, the developing verbal voice of the child, his/her social speech, will manifest dialogical positions which were "offered" to him by the caregiver(s). Thus, in speech there are certain positions related to the structure of turn-taking which are not only a matter of language convention but also socioculturally motivated.

First, the caregiver will take up all the roles and all the non-verbal as well as verbal actions, establishing in this way a model of dialogicity, speaking with more than one voice. Through this, she demonstrates the conventions of turn-taking in verbal exchanges and gives the child the opportunity to learn where and when to take up his/her role by adequate means (Stern et al, 1975; Jochens, 1979). Different voices (and positions) are marked by the role within the turn structure. Thus, a role is first bound to structure, like a scaffold to move onto, and later becomes a genuine role in terms of a position experienced as related to a certain perspective and voice. Dialogue and voice are supporting structures (Nino & Bruner, 1978). There is at first a perceived, voiced outer structure, with which the infant can concretely align. This structure then becomes

felt and experienced as an inner, meaningful one. The double-voicedness of the caregiver not only demonstrates to the infant his/her dialogical role in conversation but also the fact that a person may speak with more than one voice. And the caregiver uses fictitious voices to amuse her child who fulfils a special, playful role in extending the usage of voices and positions, demonstrating creativity and the possibilities of interacting with the world (see Josephs', 2002, case as an illustration of such an extension). This stage is termed monophonic dialogicity, for there is only the voice of the mother speaking for both the infant and herself.

As soon as the child takes up his/her role in the dialogical exchange, beginning on the verbal level with vocalizations at the right time and place, diphonic dialogicity emerges, enriching and intensifying the exchanges between mother and child. As it develops and refines, this stage leads to mature dialogues between interactants capable of attuning themselves to one another. This stage closely links the first one to the third one: the stage of polyphonic dialogicity, for it takes up the demonstrated mother's multi-voicedness from the first stage (Fogel et al, 2002) and realizes it in the ability to imagine other perspectives and enact them with their voices.

Imitation. An important device in the depicted process, especially in regard to the assumed scaffold the child can move onto, is imitation. Imitation can be seen as a means to slip into the other and his/her perspective. This ‘slip into’ is particularly interesting for it leads to an inside, rendered possible through the (as it seems) specifically human intersubjectivity. Primary intersubjectivity is related to the affective attunement of infants to their caregivers before the age of 4 months, displayed through contingent smiling, gazing, and other socially elicited gestures (see Rochat et al., 1999). Secondary intersubjectivity starts with the manifested mutual engagement as displayed in joint attention. By the age of 9 months, children begin to be able to include an object in their exchange with an adult, and they both can now behave commonly toward this object. Tomasello (1993) stresses the aspect of perspective taking which I loosely termed as ‘slip into’: “Joint attention is not just shared visual gaze but a true perspective taking.” (p.176). The ‘slip into’ is also seen in the way humans learn from each other, termed by Tomasello (1993) as cultural learning:

> In cultural learning, learners do not just direct their attention to an individual and its behavior, they actually attempt to see the world as the other individual sees it – from inside the other's perspective, as it were. It is learning in which the learner is attempting to learn not *from* another but *through* another. (p. 175, emphasis in the original)

In another study Call and Tomasello (1995) demonstrate that this form of learning is related to imitation, in contrast to what the authors call emulative learning and learning through mimicking observed in different apes. Children clearly used

information from their observations in order to solve the problem presented, imitating aspects of the demonstrated actions. Following Call and Tomasello (1995), imitation is, in contrast to emulative and mimicking learning, based on the understanding of the goals, i.e., intentions of others, the ability to understand the actions of others as goal directed. Recently, Tomasello et al (2005) have deepened this aspect on the basis of new empirical findings. Going beyond the assumption of understanding the intentional actions and perceptions of others the authors suggest "shared intentionality" as a key requisite to human cultural cognition. Thus, the ontogenetic pathway goes from dyadic engagement with shared emotions and behavior, through triadic engagement with shared goals and perceptions to collaborative engagement with joint intention and attention. Highly interesting for the theory of the dialogical self, the authors assume "a special kind of shared motivation in truly collaborative activities" (p.690), this motivation can be described as desire towards the other, as strong drive, in the end responsible for uniquely human cognition:

> Our proposal is that the uniquely human aspects of social cognition emerge only as *uniquely human social motivation* to interact with an emerging, primate-general understanding of animate and goal-directed action – which then transforms the general ape line of understanding action into the modern human line of shared intentionality. (Tomasello et al, 2005, p.688; emphasis added)

Some aspects of imitation can be added further to conceive the slip into someone's perspective. First, it should be stressed that imitation is not only quite frequent in adult-child talk, the frequency of the adult imitating the child is also worth noting (see Blount, 1972). That is, in a sole child imitation one can see dialogue which proceeds as follows:

1. child utters / vocalizes
2. adult imitates child's utterance/vocalization
3. child imitates adult's imitation of his/her own original utterance
4. adult confirms child's imitation as genuine utterance

Especially step (3) is highly interesting. The child, in imitating the adult model of his/her own first utterance, imitates or repeats him/herself but at the same time both voices are present in step (3). Of course, the utterance (or vocalization) changes in quality from (1) to (4): it is shaped according the relevance criteria valid for the specific utterance situation. This is the typically identified function of imitation: to give the child the right speech model, confirming and correcting the child's speech. Field (1978) argues in the same vein and highlights the importance of mutual imitation for dialogical development in the sense of coherence in exchange. The notion of a "mimetic spiral"

reinforces the idea of mutuality in imitation, and has the worth noting effect of changing the initial social context (Gebauer & Wulf, 1995).

Hence, the notion of imitation is a strongly dialogical one: Imitating is done in an incessant movement from one to the other, each one giving and taking parts of what is expressed, transforming it in the course of the movement. In speech acquisition, forms are established for the sake of inter-individual comprehension. It is not only this instructive function which matters, but also its bounding and carrying one which is, of course, highly affect-laden.

Vygotsky was also aware of the important role of imitation in social learning, seeing imitation as a form of collaboration between the child and others. In imitating, the child learns a new behavior, and it is the social group as psychological fact that motivates the imitations of the child. Moreover, Vygotsky depicts exactly what I try to grasp with the notion of the 'slip into' the inside of another:

> He [the child] merges in his activity with the one he imitates. The child never imitates movements of nonliving objects, for example, the swing of a pendulum. Obviously, his imitative actions arise only when there is personal communication between the infant and the person whom he imitates. (Vygotsky 1998, p. 236)

In Mauss (1936/1999) one can find the beautiful term *imitation prestigieuse*, which expresses quite well what Vygotsky describes here. Although meant in another context, the term may serve to note that one does not imitate just anybody, but a specific, "prestigious" person: a meaningful other.

In closing these remarks on the concept of imitation it is proposed that the most powerful scaffold for the child to align (and for the adult, too) is in repeating and imitating the voice quality of the other. Thus, the structure of dialogical turn-taking and of the mother's voice intonation function as supports by virtue of a concrete perceptibility (rhythm, prosody) that the infant can imitate. Children in a preverbal stage seem indeed to avail themselves of the intonation in order to come into speech and into specific speech acts like questions and demands. On this point, Bruner (1975) speaks of a "prosodic envelope":

> A fourth process [...] consists of the child learning phonological patterns almost as place-holders, imitatively. They constitute, even pre-verbally, a kind of prosodic envelope or matrix into which the child "knows" that morphemes go [...] There is the possibility that distinctive "speech acts" are learned in a primitive fashion by this means – demand prosody involving rising intonation, etc. (Bruner 1975, p. 10)

In recent researches one can find strong support for the position stressing the importance of concrete voice perception in psychological development. Castarède and Konopczynski (2005) take into account the speaking subject who has disappeared in "pure" linguistics, and to highlight with him the vocal relation between two voices. The research reported is mostly undertaken in clinical contexts from a psychoanalytic perspective, centered on the auditive-vocal exchanges between mother and very young infant (see in this line Muratori & Maestri, this issue). Noticeable also is the relation to intersubjectivity theory and to theory and research in music, both approaches being united in Trevarthen's recent research (see Trevarthen & Gratier, 2005; besides, the Special Issue of *Musicae Scientiae*, 1999-2000).

Thus, this approach supports the aspect of voice stressed here as a real auditory-vocal event right at the beginning of development. This leads to the question formulated above: How does an internal I-position develop out of external experiences with audible voices? The concept of internalization will provide a starting point in the attempt to find an answer to this complex question.

Internalization. Vygotsky (e.g., 1978) was the one who pointed to the social dimension of internalization; to view internalization as founded in social processes. Vygotsky deduced that any so-called higher (culturally determined) mental function (such as remembering, attention, thinking) develops by internalization processes out of social interactions and is thus itself fundamentally social. The interactions with others are semiotically mediated, especially by language (see Wertsch & Stone, 1985, for stressing this "semiotic mechanism" of internalization). What is internalized is the social relationship, a dynamic structure of otherness of a certain quality, mediated and at the same time shaped by language. Vygotsky does not assume that external and internal processes are copies of one another but that internalization transforms the social, inter-individual process itself and changes its structure and functions (Wertsch & Stone, 1985, p. 167). In stressing the social, inter-individual origin of individual psychological processes, Vygotsky's approach is quite close to dialogical theory: both employ a notion of alterity. For this reason, this approach will serve to develop the concept of internalization.

Keiler (2002) asserts that there are two versions of the notion of sociocultural development in Vygotsky, the first one dating from 1928 to 1930, the second one, a revision of the first, from 1931. In both versions, internalization is a key concept, whereby the role of the other is slightly changed.

Following Keiler (2002), the third of three main characteristics in the first version of Vygotsky's theory of sociocultural development is that the genesis of higher mental functions is accomplished in four stages. The first is the stage of natural psychology, followed by the stage of naive psychology, itself followed by the stage of outer cultural method with signs which are only shortlived, leading to the fourth stage

of inner activity: here internalization takes place, the outer means (signs) are transformed in inner ones, they become "ingrown": "The external means, so to speak, become ingrown or internal" (Vygotsky, 1929, p. 426). This process corresponds to a qualitative transformation of "natural psychisms" into culturally determined higher mental functions. What makes this transformational process possible is the fact that the child takes a "psychological attitude" toward him/herself, and that he/she seeks to control his/her own behavior, including mental processes (e.g., attentional and remembering processes). However, what is not mentioned here is the role of the other in forming the child's "psychological attitude", and his/her control over the child's behavior. Vygotsky explains this "ingrowing":

> Such "complete ingrowing" is based on the fact that inner stimuli are substituted for the external ones. The traces replace irritation. (Vygotsky 1929, p. 426)

Another metaphor for the transition from the third to the fourth stage is "seam-like ingrowing": connecting two organic parts together, contributing to the formation of a connecting texture and becoming itself superfluous – like the sign which becomes superfluous after some repetitions (see Vygotsky, 1929, p. 426). Vygotsky's notion so far is that of an organic process where something is growing in a certain way, backgrounding in my opinion the social aspect of internalization Vygotsky comes to underline later on.

In generalizing the four stages to any higher mental function, Vygotsky derives two main age levels where the role of the other is hinted at, the focus however remaining on the child. There is first a process from adult to child in which the child appropriates by an act of synthesizing the originally distributed process. This unified process, however, psychologically remains "distributed", in the act of *as if* it was done by two persons, relating these persons and their behavior. Doing this, the child is then able to "grow in", that is, to move from outer use of means to an inner one. I think there is here a structural similarity to the process of imitation as sketched above. In imitation, too, the child takes a behavior from the adult and performs it *as* its own. This is possible because the imitated behavior was originally own's one, imitated by an adult. So, imitating has a double-voicedness, and internalization too, for it brings together other and self in one person.[9]

The revised version is dated by Keiler (2002) autumn 1930 with Vygotsky's conference on psychological systems (Vygotsky, 1997a). The development of higher mental functions is now not limited to purely intrafunctional change, where any function is transformed as such, but to a deep interfunctional one, where the original relations between functions are transformed. From this, Vygotsky derives his well-

[9] See Vygotsky's work on Pedology, 1931, in Vygotsky, 1998, p. 104.

known claim that any higher mental function appears twice in development: first in a collective form as interpsychic function, second in an individual form as intrapsychic function (see Vygotsky, 1978, p. 57). This leads to the general thesis that any higher mental function is originally a reciprocal, mutual process. Having described the child as taking a "psychological attitude toward his/herself" in his first version of sociocultural development, Vygotsky speaks in late 1930 of the child taking the role of the mother toward his/herself, and coming, through this role reversal, to control his/her own behavior (Keiler, 2002, p. 197). So, the turning around, so to speak, is now given a clear social origin.

In his 1931 published work on "Pedology" (Vygotsky, 1998), Vygotsky heightens the social aspect in his second law on the cultural development of behavior, saying that the relations between higher mental functions are transferred social relations. So, the higher functions are inner social relations, transferred into the personality, but deeply social. Language is in this transfer of critical importance, understood by Vygotsky as a means of influence, of acting on another and – by internalization – acting on oneself (Keiler, 2002, p. 201). This is what permits Vygotsky to write: "Through others, we become ourselves" (1998, p. 170); the development of the child involves the transformation of social relationships in mental functions (Vygotsky, 1997b).

Even though Vygotsky quite precisely depicts the aspects of internalization through the metaphors of "ingrowing" and "seam-like ingrowing" and in the processes of "as if" and role change beween mother and child, whereby language leads and shapes the process, I would like to try to come even a bit closer to the mechanism of internalization, stated by Wertsch and Stone (1985) as "semiotic". I propose to narrow "semiotic" to verbal signs, thus following Vygotsky's acknowledgement of language as a privileged means of internalization – being himself in the tradition of Hegel's view that language is the tool of tools (see Keiler, 2002, p. 188). So, the point would be to come closer to the language process.

It is already clear that imitation and internalization are closely related. Both function through an 'other' with whom self is acting. This other is – besides being a prestigious one to whom the self wishes to be related – physically present, a mirror and a former of the self's actions and expressions. The presence of the other is thus active, not just there but directed, addressed to self, literally in touch with self: by means of actual touch or of voice, or of both. And this is done in physical, reiterated patterns: forms giving themselves form in the growing mutuality of an adult and an infant (Lyra, 2007, this issue). Imitation allows an exchange of forms of behavior and forms of expression, an exchange corresponding to a close give and take. Role change distinguishes more clearly between what the one and the other is doing, and allows one, therefore, to be the other for a moment, to integrate this other in self. Again, voiced forms play an important role, giving the child indications about roles and their timing.

Imitation, role change and as-if acting, which are all found in children's play and in their imaginative dialogues, are devices in the process of moving from the outside to the inside, and between self and other. Lillard (2001) points to the fact that in pretend role play and in imaginary companion pretence "a child practices at being other people" and comes to experience and thereby know the other's thought – thus, pretence is important to theory of mind. Further, in pretend play children construct a "decoupled world", an operation by which representations are temporarily removed from their usual referents, also described as "conceptual move". I suggest that this move is akin to the one I posit, leading to a reconstructed and transformed other in self, still speaking, a resounding trace in memory and imagination. The basis of the move is a sensitive experience of the other, and this is the reason for giving the actual voice, in which language is expressed and given to another, a specific status: it is a live form which acts as a carrier leading from outside to inside. This form is form and meaning, or: formed meaning. Meaning which manifests itself in form, not detachable from it. Noteworthy for the process of moving inside, a precise (or adult) meaning need not be established for the form to function as carrier. Indeed, Vygotsky (1987) underlines the contrast between adult's and children's concepts, different in meaning but seemingly the same because of their linguistic form. The child's concept, the inner side of meaning is developing. As Wertsch and Stone (1985) formulate:

> One of the correlates of the fact that interpsychological semiotic processes requires the use of external sign forms is that it is possible to produce such forms without recognizing the full significance that is normally attached to them by others. As a result, it is possible for a child to produce seemingly appropriate communication behavior before recognizing all aspects of its significance as understood by more experienced members of the culture. One of the mechanisms that makes possible the cognitive development and general acculturation of the child is the process of coming to recognize the significance of the external sign forms that he or she has already being using in social interaction. (p. 167)

What I suggest is to see the experienced voice of a significant other as mechanism of internalization. The specific intonations and the expressive, idiosyncratic style of the person as manifested in her voice give a specific taste as to what is internalized: this is individual as well as inter-individual, corresponding to the genres of speaking and intonating of the speech community. So, what permits the movement from outside to inside is a meaningful, perceivable social form, tied to a person. I understand the voice as this form, carrying the other into self and self into other, a scaffold: graspable, embodied and thus living materiality. This form offers a meaningful structure in so far as it is always turned toward somebody and because of its appertaining to the inter-individual interactional world it is rooted in. Both ways of having and giving

meaning – in personally addressing and in being inter-individually rooted - are indissociable, assuming that the individual alone is non-existent, solely conceivable as a social being whose psyche and consciousness are socio-ideological facts (Voloshinov, 1973, p.12, 34). As Voloshinov (1973, p. 22) writes, for the animal cry "the social atmosphere is irrelevant", this cry is bereft of any value accent. But a voice does count on such an atmosphere and it sets an ideological accent, thus belonging to the inter-individual realm.

In the course of development the voice as perceivable form is interiorized, and with it the attitude and perspective of the (social) person the voice belongs to. In the dialogical self, several voices exist on the basis of primary voice experience. Some may retain their relation to a specific person, some may be altered by such processes as condensation and displacement, by imagination and generalization (see Mead's generalized other). A completely abstracted voice is then conceivable as a subject's perspective and conceptual horizon – but the primary experience bound to the perception of a speaking person as present body is the necessary ground.

Conclusion

In my reading of Bakhtin, a voice has the function of a carrier as it carries the speaking subject out of himself, decentering and orienting him toward the other(s), supporting and leading the contact. What a voice carries and expresses at the same time is that the utterance is as well "mine" as "other's" (Bakhtin, 1986, p.89). A voice carries the individual expression of contact with the other which is always mingled with some alien components. It supports the necessary multiplicity belonging to the living language in the word.

It is this idea of a carrier that serves to understand the phenomenon of voice, coupled with the Russian ideas of life or vividness and of the materiality of the verbal sign, existing in concrete verbal interactions (Voloshinov, 1973). Therefore, my own accent in the notion of voice is on its experienced and embodied, material and social dimensions. In this, I meet Osatuke's et al (2004) notion of the "physicality of the psychological self". Setting the point of departure in ontogenesis, it is the resounding quality of voice that matters, its *experienced form*: thus intonation, understood according to Bakhtin and Voloshinov as belonging to the social and as manifesting ideology, becomes quite a central feature of the notion of voice developed here. The features of indexicality and body stress the participation of voice in the social as well as in the individually perceived and experienced world of humans. With the concepts of imitation and internalization I tried to explain how this experienced form is transformed into a cognitive perspective. In this context, it is central that voice belongs to a fundamental structure of addressivity, and exists only within this structure, expressing and shaping it as well. I call the voice a form which is form and meaning, to be

understood from perceptual and interactional experience. Thus, I am oriented towards Bühler's notion of form:

> the form of something is there to the end of pointing out an idiosyncrasy in which the form is realized. (Bühler 1934/1990, p. 129)

In internalization, this form shows its power. The experienced voice of a significant other is the mechanism of internalization. What permits the movement from outside to inside is a meaningful, perceivable social form, tied to a person. The voice is a form of vivid materiality, it offers a meaningful structure in so far as it is always turned toward somebody. And it is meaningful because of its participation in the inter-individual interactional world it is rooted in.

The idea of language related to this notion of voice is that it is not only and not foremost an abstract system involving elements, rules, and concepts, but that it is first a perceivable event between persons, performed by these persons on the foundations of their structure of addressivity belonging to them as human beings. This event takes place as form, rooted in sensory experience taking place in interactions, thus, developing its specific verbal and voiced form on the basis of preverbal (vocalizations, sounds) and non-verbal forms (rhythms, routines, patterns) of mutuality. The root in sensory experience links language on the one hand to the body, and on the other hand to a physical environment others are the most important part of. For this reason, it is essential that what serves language processes is embodied. Our privileged embodiment is voice: an auditory-vocal event, belonging to the realm of experience, both in self and in other. And able to be detached from this realm in order to enter abstract, symbolic meaning serving intra-mental processes. To understand how language functions, to understand its specificity as a linguistic system, I believe one has to go beyond it – in the way the Russian thinkers, Yakubinsky and Voloshinov, have shown. And the notion of voice as developed here serves this goal, linking language back to our body and to the others.

Acknowledgements

I wish to express my gratitude to the friends and colleagues from what I would call the *Cercle de Crêt-Bérard* (CH) who invited me repeatedly to share their discussions about Russian-Soviet thinking on language from the perspectives of (psycho-)linguistics, philosophy, and Slavistics, thus putting me on the trail: Cristian Bota, Janette Friedrich, Irina Ivanova, Patrick Sériot and Kim Stroumza. Finally, I am grateful to Miguel Gonçalves and to Peter Raggatt for their precious comments, as well as to Giancarlo Dimaggio for his constructive critical remarks on previous drafts of this article.

References

Akhtar, N. & Tomasello, M. (1998). Intersubjectivity in early language learning and use. In S. Bråten (Ed.), *Intersubjective communication and emotion in early ontogeny* (pp. 316-335). Cambridge, UK: Cambridge University Press.

Bachtin, M. M. (1979). *Die Ästhetik des Wortes* (R. Grübel, Ed.). Frankfurt am Main: Suhrkamp.

Bachtin, M. M. (2004). Das Problem der sprachlichen Gattungen (J. Dittmar & K. Meng, Trans.). In K. Ehlich & K. Meng (Eds.), *Die Aktualität des Verdrängten. Studien zur Geschichte der Sprachwissenschaft im 20. Jahrhundert* (pp. 447-484). Heidelberg: Synchron.

Bakhtin, M. M. (1984). *Problems of Dostoevsky's poetics* (C. Emerson, Ed., & Trans.). Minneapolis: University of Minnesota Press.

Bakhtin, M. M. (1986). *Speech genres and other late essays* (C. Emerson & M. Holquist, Eds., V. W. McGee, Trans.). Austin: University of Texas Press.

Bertau, M.-C. (1999). Spuren des Gesprächs in innerer Sprache. Versuch einer Analyse der dialogischen Anteile des lauten Denkens. *Sprache & Kognition, 18*, 4-19.

Bertau, M.-C. (2004a). Developmental origins of the dialogical self: Some significant moments. In H. J. M. Hermans & G. Dimaggio (Eds.), *The dialogical self in psychotherapy* (pp. 29-42). New York: Brunner-Routledge.

Bertau, M.-C. (2004b). Reflections on addressivity: From the role of the other to developmental aspects. In M.-C. Bertau (Ed.), *Aspects of the dialogical self* (pp. 87-128). Berlin: Lehmanns Media LOB-de.

Bertau, M.C. (2004c). The theory of the dialogical self and a proposition for modeling. In M.-C. Bertau (Ed.), *Aspects of the dialogical self* (pp. 9-35). Berlin: Lehmanns Media LOB-de.

Bertau, M.-C. (2007, in press). Pour une notion de la forme linguistique comme forme vécue. In P. Sériot & J. Friedrich (Eds.), *Recueil Jakubinskij – Voloshinov – Vygotski*. Lausanne.

Blount, B. G. (1972). Parental speech and language acquisition: Some Luo and Samoan examples. *Anthropological Linguistics, 14*, 119-130.

Bruner, J. S. (1975). The ontogenesis of speech acts. *Journal of Child Language, 2*, 1-19.

Bruner, J. S. (1983). *Child's talk: Learning to use language*. London: Norton.

Bühler, K. (1990). *Theory of language. The representational function of language* (D. F. Goodwin, Trans.). Amsterdam, The Netherlands: John Benjamins. (Original work published 1934)

Call, J. & Tomasello, M. (1995). Use of social information in the problem solving of orangutans (*pongo pygmeus*) and human children (*homo sapiens*). *Journal of Comparative Psychology, 109*, 308-320.

Castarède, M. F. & Konopczynski, G. (Eds.) (2005). *Au commencement était la voix.* Ramonville Saint-Agne: Érès.

Coulet, C. (1996). *Communiquer en Grèce ancienne. Écrits, discours, voyages.* Paris: Les Belles Lettres.

Dimaggio, G. & Hermans, H. J. M. (2006, submitted). *Dialogical self theory: New formulation and implications for research and practice*.

Ehlich, K. & Meng, K. (Eds.) (2004). *Die Aktualität des Verdrängten. Studien zur Geschichte der Sprachwissenschaft im 20. Jahrhundert*. Heidelberg: Synchron.

Field, T. (1978). The three Rs of infant-adult interactions: Rhythms, repertoires, and responsivity. *Journal of Pediatric Psychology, 3*, 131-136.

Fogel, A., de Koeyer, I., Bellagamba, F., & Bell, H. (2002). The dialogical self in the first two years of life. *Theory & Psychology, 12*, 191-205.

Gasparov, M. (1982). Michail Bachtins Stellung in der russichen Kultur des 20. Jahrhunderts. In R. Lachmann (Ed.), *Dialogizität* (pp. 256-259). München: Wilhelm Fink.

Gebauer, C. & Wulf, C. (1995). Social mimesis. *Paragrana*, *4*, 13-24.

Hanks, W. F. (2001). Indexicality. In A. Duranti (Ed.), *Key terms in language and culture* (pp. 119-121). Malden, MA: Blackwell.

Hermans, H. J. M. & Hermans-Jansen, E. (1995). *Self-narratives: The construction of meaning in psychotherapy*. New York: Guilford.

Hermans, H. J. M. (1996). Voicing the self: From information processing to dialogical interchange. *Psychological Bulletin, 119*, 31-50.

Hermans, H. J. M., Kempen, H. J. G., & van Loon, R. J. P. (1992). The dialogical self. Beyond individualism and rationalism. *American Psychologist, 47*, 23-33.

Hermans, H. J. M & Kempen, H. J. G. (1993). *The dialogical self: Meaning as movement*. San Diego, CA: Academic Press.

Hermans, H. J. M. & Dimaggio, G. (2004). *The dialogical self in psychotherapy*. New York: Brunner-Routledge.

Ivanova, I. (2003). Le dialogue dans la linguistique soviétique des années 1920-1930. *Cahiers de l'ILSL ,14*, 157-182.

Jakubinskij, L. (2004). Über die dialogische Rede. In K. Ehlich & K. Meng (Eds.), *Die Aktualität des Verdrängten. Studien zur Geschichte der Sprachwissenschaft im 20. Jahrhundert* (pp. 383-433). Heidelberg: Synchron. (Original work published 1923) [Note: Some of the core passages of Yakubinsky's article paraphrased in

English from the German translation are available under www.bertau.de/en/index.html in the “Archives”]

Jochens, B. (1979). “Fragen” im Mutter-Kind-Dialog: Zur Strategie der Gesprächsorganisation von Müttern. In K. Martens (Ed.), *Kindliche Kommunikation. Theoretische Perspektiven, empirische Analysen, methodische Grundlagen* (pp.110-132). Frankfurt am Main: Suhrkamp.

Josephs, I. (2002). “The Hopi in Me”. The construction of a voice in the dialogical self from a cultural psychological perspective. *Theory & Psychology*, *12*, 162-173.

Karsten, A. (in preparation). *Dialoges Schreiben. Relevante Basisprozesse*. Magisterarbeit, Universität München.

Keane, W. (2001). Voice. In A. Duranti (Ed.), *Key terms in language and culture* (pp. 268-271). Malden, MA: Blackwell.

Keiler, P. (2002). *Lev Vygotskij – ein Leben für die Psychologie*. Weinheim und Basel: Beltz.

Laver, J. (1975). Communicative functions of phatic communion. In A. Kendon, R. M. Harris, & M. Ritchie de Key (Eds.), *The organization of behavior in face to face interaction* (pp. 215-237). The Hague: Mouton.

Lillard, A. P. (2001). Pretend play as twin earth: A social-cognitive analysis. *Developmental Review*, *21*, 495-531.

Lyra, M. D. C. P. (2007). On abbreviation: Dialogue in early life. *International Journal for Dialogical Science*, *2*, 15-44. Available at http://ijds.lemoyne.edu/journal/2_1/IJDS.2.1.02.Lyra.html

Mauss, M. (1999). Les techniques du corps. In M. Mauss, *Sociologie et anthropologie* (pp. 365-386). Paris: Quadrige/Presses Universitaires de France. (Original work published 1936)

Mead, G. H. (1967). *Mind, self, and society from the standpoint of a social behaviorist*. Chicago: The University of Chicago Press (Original work published 1934)

Mehrabian, A. (1972). *Nonverbal communication*. Chicago: Aldine-Atherton.

Meng, K. (2004). Das Konzept der Äußerung bei Bachtin und Vološinov. In K. Ehlich & K. Meng (Eds.), *Die Aktualität des Verdrängten. Studien zur Geschichte der Sprachwissenschaft im 20. Jahrhundert* (pp. 153-190). Heidelberg: Synchron.

Musicae Scientiae. (1999-2000). Special issue: *Rhythms, musical narrative, and the origins of human communication*. Liège: European Society for the Cognitive Sciences of Music.

Naumova, T. (2004). Das Problem des Dialogs: A. A. Potebnja, L. P. Jakubinskij, L. S. Vygotskij, M. M. Bachtin. In K. Ehlich & K. Meng (Eds.), *Die Aktualität des*

Verdrängten. Studien zur Geschichte der Sprachwissenschaft im 20. Jahrhundert (pp. 211-225). Heidelberg: Synchron.

Ninio, A., & Bruner, J. (1978). The achievement and antecedents of labelling. *Journal of Child Language, 5*, 1-15.

O'Connell, D. C., & Kowal, S. (2003). Psycholinguistics: A half century of monologism. *American Journal of Psychology*, *116*, 191-212.

Osatuke, K., Gray, M. A., Glick, M., Stiles, W. B., & Barkham, M. (2004). Hearing voices. Methodological issues in measuring internal multiplicity. In H. J. M. Hermans & G. Dimaggio (Eds.), *The dialogical self in psychotherapy* (pp.237-254). New York: Brunner-Routledge.

Osatuke, K., Humphreys, C. L., Glick, M., Graff-Reed, R. L., McKenzie Mack, L., & Stiles, W. B. (2005). Vocal manifestations of internal multiplicity: Mary's voices. *Psychology and Psychotherapy: Theory, Research and Practice, 78*, 21-44.

Radzikhovskii, L. A. (1986-87). The dialogic quality of consciousness in the works of M. M. Bakhtin. *Soviet Psychology, 25*, 3-28.

Raggatt, P. (2006). Multiplicity and conflict in the dialogical self: A life-narrative approach. In D. P. McAdams, R. Josselson, & A. Lieblich (Eds.), *Identity and story: Creating self in narrative* (pp.15-35). Washington, DC: APA Press.

Romashko, S. A. (2000). Vers l'analyse du dialogue en Russie. *Histoire, Épistémologie, Langage*, *22*, 83-98.

Rochat, P., Querido, J. G., & Striano, T. (1999). Emerging sensitivity to the timing and structure of protoconversation in early infancy. *Developmental Psychology*, *35*, 950-957.

Saenger, P. (1989). Psychologie de la lecture et séparation des mots. *Annales (Économies, Sociétés, Civilisations), 44*, 939-952.

Schütz, A. (1982). *Das Problem der Relevanz*. Frankfurt am Main: Suhrkamp.

Snow, C. E. (1977). Mother's speech research: From input to interaction. In C. E. Snow & C. A. Ferguson (Eds.), *Talking to children* (pp. 31-49). Cambridge, UK: Cambridge University Press.

Stiles, W. B. (1999). Signs and voices in psychotherapy. *Psychotherapy Research 9*, 1-21.

Stiles, W. B., Osatuke, K., Glick, M., & Mackay, H. C. (2004). Encounters between internal voices generate emotion. An elaboration of the assimilation model. In H. J. M. Hermans & G. Dimaggio (Eds.), *The dialogical self in psychotherapy* (pp. 91-107). New York: Brunner-Routledge.

Stern, D. N., Jaffe, J., Beebe, B., & Bennett, S. L. (1975). Vocalizing in unison and in alternation: Two modes of communication within the mother-infant dyad. *Developmental psycholinguistics and communication disorders. Annals of the New York Academy of Sciences, 263,* 89-100.

Tannen, D. (1989). *Talking voices: Repetition, dialogue, and imagery in conversational discourse*. New York: Cambridge University Press.

Tomasello, M. (1993). On the interpersonal origin of self-concept. In U. Neisser (Ed.), *The perceived self: Ecological and interpersonal sources of self-knowledge* (pp. 174-184). Cambridge, UK: Cambridge University Press.

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Brain and Behavioral Sciences, 28*, 675-691.

Trevarthen, C., & Gratier, M. (2005). Voix et musicalité: Nature, émotion, relation et culture. In M. F. Castarède & G. Konopczynski (Eds.), *Au commencement était la voix* (pp. 105-116). Ramonville Saint-Agne: Érès.

Voloshinov, V. N. (1973). *Marxism and the philosophy of language*. London: Seminar Press. (Original work published 1929)

Voloshinov, V. N. (1981a). Le discours dans la vie et le discours dans la poésie. In T. Todorov, *Mikhaïl Bakhtine, le principe dialogique, suivi de Écrits du Cercle de Bakhtine* (pp. 181-215). Paris: Seuil. (Original work published 1926)

Voloshinov, V. N. (1981b). La structure de l'énoncé. In T. Todorov, *Mikhaïl Bakhtine, le principe dialogique, suivi de Écrits du Cercle de Bakhtine* (pp. 287-316). Paris: Seuil. (Original work published 1930)

Vygotsky, L. S. (1929). The problem of the cultural development of the child. *Journal of Genetic Psychology, 36*, 415-434.

Vygotsky, L. S. (1987). *Thinking and speech* (N. Minick, Ed. & Trans.). London: Plenum. (Original work published 1934)

Vygotsky, L. S. (1978). *Mind in society*. *The development of higher psychological processes* (M. Cole, V. John-Steiner, S. Scribner, & E. Souberman, Eds.). Cambridge, MA: Harvard University Press.

Vygotsky, L. S. (1997a). *Problems of the theory and history of psychology* (The collected works of L. S. Vygotsky, Vol. 3). New York: Plenum Press.

Vygotsky, L. S. (1997b). *The history and the development of higher mental functions* (The collected works of L. S. Vygotsky, Vol. 4). New York: Plenum Press.

Vygotsky, L. S. (1998). *Child psychology* (The collected works of L. S. Vygotsky, Vol. 5). New York: Plenum Press.

Wertsch, J. V., & Stone, C. A. (1985). The concept of internalization in Vygotsky's account of the genesis of higher mental functions. In J. V. Wertsch (Ed.), *Culture, communication, and cognition: Vygotskian perspectives.* (pp. 162-179). Cambridge, UK: Cambridge University Press.

(This page intentionally left blank)