

Antje Vowinckel  
TERRA PROSODIA – Sound compositions with dialect melodies.  
Translated by Kate Donovan

Speech melodies: at home they are invisible or drab, but in foreign lands they whirl their way around like colourful butterflies. And are just as difficult to catch. Approach them with a microphone, as with a net, and they quickly disappear. Not the people, but the melodies. If you do manage to catch one, it often seems the net becomes a steamroller, flattening out the melody – the sight of a microphone makes most people immediately speak more flatly.

My interest in dialect melodies<sup>1</sup> is not so much concerned with regions and traditions, but rather the unintentional, fleeting music therein. I first found such unintentional melodies in international language courses – right there where regional language differences are supposed to be overcome. For my piece *Call Me Yesterday*, I collected audio language courses and came across distinct modulations throughout, which sounded like artificial melodies but were never intended as music. The melodies were created solely out of the will to perfect pronunciation, and only emerged as a "side effect" of the didactic work. They are objets trouvés, so to speak, which can only act musically once in a different context. To find them, however, you have to listen around a lot of meaning, and through many consonants. Thus confirming Tom Johnson's observation, "it often takes longer to find a good piece than to compose one."

It was this search for such accidental melodies that later brought me to dialects. But first, in composing with speech melodies, there are many different forerunners.

---

### Forerunners

Even in the ancient world, melody and prosody parted ways. To melody were attributed harmony, metrics and instrumentation (song); to prosody, speech – a single term, encompassing not only speech-melody, but also tempo, rhythm, volume, accent, and pauses. Melody was seen as something figurative and clearly recognisable, added to language. Speech-melody, in contrast, was by definition "hidden" within prosody, bound within a bundle of parameters. The control of this served primarily in making the meaning of the words vividly understood.<sup>2</sup>

As speech-melodies emerged in the European tradition around 1600 as recitative, these served primarily to convey more content and to advance the opera's plot. In a rather artificial way, it was "naturally" spoken; melodies were devised and music was adapted somewhat to the rhythm of speech. Not until Schönberg's *Pierrot Lunaire* (1912), in which the composer notated a melody that should neither be sung nor spoken, was it clearly more about the musical expression of speech than the content. This musical expression of speech also inspired Partch in his compositions for intoning voice. In his *Seventeen Lyrics for Li Po* (1931-33) and subsequent works, Partch conveyed the microtonal modulation of speaking in the accompanying instrumental voices. While he went to

---

<sup>1</sup> I do not make a scholarly differentiation between dialects and minority languages. For my recordings, I am looking for speech melodies that are as distinctive as possible. Most often they are found in the dialects of isolated regions, but the English in Glasgow also has characteristic melodies, for example.

<sup>2</sup> Hans Günther Tillmann helpfully differentiates between A-, B- and C- prosody, whereby the A-prosody can be controlled, the B-prosody occurs spontaneously. Hans Günther Tillmann, Phil Mansell: *Phonetik. Lautsprachliche Zeichen, Sprachsignale und lautsprachlicher Kommunikationsprozess*. Stuttgart, Klett-Cotta 1980.

great lengths and invented many special instruments to reflect these intonations – overheard from people in the street – his intoning voice parts (although rhythmically free and alternating between speaking and singing voices) change exactly this natural intonation by artificially declaiming the texts. In this way spontaneous modulation (or B-prosody, according to Tillmann) was remodulated through controlled declamation (or A-prosody). In some short passages of *Bitter Music* (1935), the instrumental part seems to completely follow the impulsive intonations of speech, which has to do with the fact that the speaker speaks unaccompanied for a long time. However, the speaking voice is also written out; the speaker reads the text. It remains – even if diminished - a declamatory modulation. This would be taken up by the piano, which often then transformed it into song melodies.

The possibility of audio recording on the one hand, and the opening of the arts toward found material on the other, has accelerated the awareness of spontaneous, natural speech as music since the 1930s. Walter Ruttmann was the first to use recorded voice fragments in his composition produced on film, *Weekend* (1930). In the 1950s John Cage (*Williams mix*), as well as Pierre Schaeffer and Pierre Henry (*Symphonie pour un homme seul*), composed with tape recordings. They cut, combined and repeated language fragments, assembling such sounds into various musical structures. But even these voice recordings have, for the most part, a declamatory, artificial modulation. (Just as many other pieces composed with language, by Raoul Hausmann, Kurt Schwitters or Ernst Toch, for example). Olivier Messiaen's compositions based on birdsong do not refer to human speech, but use found melodies as a starting point. There were experiments with so-called "Original-Ton" – free, spontaneous speech – in Neues Hörspiel since the end of the 1960s in Germany; in Kagel's *Hörspiel - ein Aufnahmезustand*, for example. In these works, specific speech melodies play a somewhat less central role, but the freely improvised character of speaking can be recognised, and it is no longer completely overlaid with foreign musical forms. Natural speech became material for an artistic starting point. This also applies to Robert Ashley's work *Automatic Writing* (1979), which uses impulsively spoken phrases derived from his own experience of Tourette's Syndrome. (In other works, Ashley often rather seeks to transfer natural inflections into a music based on declamatory speaking.) Steve Reich used short fragments of railway announcements in *Different Trains* (1988), in which the turn to unformed everyday speech can also be detected. However, these announcements were not entirely spontaneous utterances. Reich lets the instruments imitate the melodies contained in these small fragments, and merge into the overlying musical structure.

Eventually, in 1989, the Canadian composer René Lussier used longer passages with spontaneous speech melodies in his work *Le Tresor de la Langue*. Here, natural speech melodies actually dominate, and are themselves imitated by instruments. They consistently return in the multi-part work, and only partially go into larger musical forms. Often, the instrumental melodies begin beforehand and lay themselves down, as if by chance, on the speech. However, Lussier scarcely trusted the perception of melody in spoken language without this doubled accompaniment. The piece has a large instrumental orchestration, it has a strong dimension of content (it is about the province of Québec), in parts it is also "scenic," and (at least in his first performance to a French-speaking audience) the audience could understand the language. It is therefore evident that speech without imitative melody would immediately be perceived as sound or content once again, rather than melody. This perception of speech is so normal for us that it can be overturned at any time. Conversely, it is easy to perceive solo instrumental passages as "chatter" when they imitate the interval sequences of speech. In this case, neither sound nor content can distract from the melody. It's all a question of focus.

Based on the ancient Greek definition, one could say that speech melody in music – as in the arts in general – is undergoing a slow paradigm shift. Step by step, it is changing from telling (diegesis) to showing (mimesis).

---

## Terra Prosodia

In my piece *Terra Prosodia* (2011), I wanted to establish a mode of perception that gradually allowed the speaking voice to be heard as an independent melody voice, without an instrument permanently emphasizing that melody. This called for a strong focus on the internal linear movement. As soon as many different instruments, sounds, or spatial depths supervene, speech will be heard externally again. You perceive the overall sound with rhythm and consonants. You visualize the person who is speaking – their age, their gender, their milieu and their region.

In dialects there are distinct speech melodies that no one intended as music. They are the result of a unique combination of spontaneous impulse and thousands of years of tradition. The melodies are tightly bound to the geographical region, to the landscape, so that you might get the impression that the valleys and mountains are continuing on in the speaking. These melodies may seem inconspicuous to their speakers, but for visitors from other regions they display a characteristic musical charm. The fact that only a small community of people speak the dialect has the advantage, in my work, that the vast majority do not understand it.

---

## Existing Recordings

The unintended music created as a by-product of languages and dialects is in immense global abundance. Linguists, however, have little interest in it. Modulation only plays a role in linguistics when it carries meaning – as, in particular, in the many tonal languages of Asia, or sometimes in the distinction between statement and question. This is evident in most historical recordings made by linguists, ever since the inception of recording equipment. In these recordings, language is almost always dissected before it actually begins to sound. Texts were read out, vocabulary lists spoken, or individual words enunciated upon request. It never ceases to amaze me that there are practically no recordings that represent the natural flow of speech. And when they actually can be found, the researcher has held the microphone too far away, or music plays in the background, or the air conditioning drones, kids play, phones ring, clocks tick, refrigerators sigh. Again and again, I thought I was on the verge of a good recording because the description sounded so promising. Nowadays there are also many dialect recordings on YouTube. They are always about learning, however, and not about native speakers. The actual dialect speakers have little interest in presenting their way of speaking on YouTube. After a while my impression was set, that dialect recordings can be compared to granola bars - they rarely contain what they claim to.

I quickly abandoned the attempt to record dialect speakers in Berlin. Anyone who has lived in a big city for a while has a flattened modulation. The impulsive, spontaneous speech disappears. You start to control yourself, to deliberate and speak more moderately. Of course you can also cultivate a pronounced modulation. Actors learn this in training. But there is a big difference between artificial modulation, which is deployed to underscore the intended meaning or emotion of a sentence, and a spontaneous modulation, which is rooted in the region (see Tillmann). Ultimately, these distinctions are theoretical, of course. Everyone is also an actor, and it's hard to make a clear distinction between personal, possibly feigned modulation and local colour. But at least I could do something to encourage them to speak in a manner that was as natural as possible.

---

## My own recordings

Because I wasn't convinced by either the existing recordings, or the dialect speakers living in Berlin, I decided to record people where they live. The seemingly simple request "tell me a story from your life," quickly turned out to be rather difficult. For most speakers living in remote areas, it is completely incomprehensible that I would travel so far just to hear any old story. They doubt that the sound of their speech could be reason enough, and so they look for a reason elsewhere – namely in their region and history. Instead of telling something personal, many people come with books, songs, or poems. Then they want to tell me the history of their village. In the next attempt, they tell their own life story – including dates. In all such accounts, the voice is flatter because the people are not emotionally involved. I really have to insist that I only want to hear the story of a single event, something that could start with "One day...," a story filled with details, and it must be something that they have experienced themselves. You need a lot of patience, and the best possible mediator, in order to get to this point. Subsequently, I let the mediator just explain the rough content of the story to me and I ask about the meaning of the most common narrative phrases like "and then..." "later..." "you know..." "that's how it was..." etc., which are always a bit varied, but they occur in every story. I can then use these as building blocks, and also repeat or shift them. Otherwise, the order of the story remains intact.

Sometimes I would get people who had developed a certain routine in storytelling, and who were well known in the village as good storytellers. That can be positive, but it can also lead to very natural modulation being covered up by narrative gesture – such as you might use for children when you want them to be lulled by a fairy tale. Instead of local colour, you hear a rolling singsong. Therefore, I do not use such well-rehearsed or memorized stories.

---

## Compositional technique

For the compositional process, I extract a MIDI-file from the audio files, thereby reducing the sound to pure melody. I consciously keep to a tempered chromatic scale, a decision which is regularly met with protest, as speech, of course, does not follow chromatic steps (as Partch already knew). And nowadays so much more is possible. But even this chromatic result already sounds so overly complex that you get lost in it, like our eyes in a cave of stalactites and stalagmites. It has little musically in common with the catchy little sentences that get stuck in your head when you hear the dialects.

My conclusion was, therefore, that it is futile to try to represent the nature of the melodies as realistically as possible. Instead, it is a matter of representing our *perception* of (rather than the *nature* of) the melodies, which is in itself influenced by familiar patterns, among them the equal-tempered tuning of the chromatic scale. In our recollection, we by no means hear the nuances that a MIDI-file or a microtonal scale reflects; on the contrary, we simplify even further. And only in this simplified mode of recollection are the typical traits of the dialects conveyed. In an accurately rendered MIDI-file (even with a sophisticated scale), you can no longer recognize the different characteristics. It beeps – flatly – in Morse-code.

I decided, therefore, to go in exactly the opposite direction and to simplify the MIDI-files further. Here, I follow my own recollection and perception. Which one is the main thread, what gets stuck? The slurring of speech I achieved by means of pitch bending the MIDI-scale, but it also shows here that too much accuracy is counterproductive. The glissandi can thereby move so intensely into the foreground that, again, the result does not correspond with our impressions of speech. In reality, we often fail to hear such details. (For sure there are experts who can also memorize microtonal intervals, and of course this perception is also dependent on the tonal system with which you grow up; different cultures have different patterns, but they could only be represented then in that particular culture.) So I used the bending flexibly, sometimes more, sometimes less, following my own perception. I then recorded the resulting melodies, either with voice, or with acoustic or virtual instruments.

At the beginning of a piece, the instrumental/singing voice mimics the speaking voice. In this way, the focus of perception is put on the melody, and the fact that it deals with a foreign language helps as well. From this point on, I trust the perception of the listener, who is handed the job of resetting their mode of reception from diegesis to mimesis. After a while – if you can hear the speaking voice as melody - the instrumental part departs from it and goes its own way, while retaining the characteristic intervals of the respective dialect. In this way, the linear character of the composition is somewhat loosened, you follow the narrative less and rather a general intervallic shape instead.

I could not depart further from this linearity, however. Any expansion into more spatial depth, the addition of other tone colours or a focus on the tone colouration of the speaking voice with its consonants – all of that distracts from the linear speech melody. Speech is once again heard externally – other prosodic parameters come to mind, you take the voice as tone colour, or even as a person, the instrumental voice appears as a foreign body again and not as a second voice.

---

## Outlook

At the moment, there are still around 6000 languages in the world, so there is a lot to do. and I still want to let other speech and dialect melodies be heard. My intention, however, is not to save these languages and dialects. In any case, only the speakers – if they like – and their countries can enable this. What I want to reveal is an impression of the tight fusion of landscape and speech, and the strong melodic potential that can be heard in it. It is a human expressiveness, which lies in traditional and, at the same time, spontaneous, impulsive speech. Eventually, as we notice that we don't hear these melodies anywhere any more, we will also notice that we have lost this specific ability to express ourselves (because we increasingly live in cities and on the Internet, and tend to deliberate before speaking). This unmediated expressive power – in addition to the desire for identity – is an important reason for dialects to become popular again. And there's no wonder it's happening up in the aeroplane of all places – the non-place of globalization par excellence. For example, when Air Berlin encourage their flight crew to make their announcements in a dialect, or like the Swabian stewardess who became known via Youtube because she just babbled away. Without fastening the vocal cords...