

Chapter 6

Summary, Conclusions, and Future directions

6.1. Summary and Conclusions

In this thesis I tackled some phonological issues from a musical perspective. In the first chapter we showed that phonological structure has a lot in common with musical structure and that both kinds of structures can be easily described using the method of Optimality Theory. Some of the similarities between speech and music have a psychological or possibly a neurological basis. In both disciplines preference rules for ideal outputs indicate the prominent constituents of every part of the hierarchical structure, and structurally important elements are assigned salience by aligning tonally or durationally strong elements with them. These resemblances led us to the assumption that insights of music theory can help out in phonological issues. Three of such issues were subjects of the experiments in this dissertation.

The first issue was the question whether the influence of a higher speaking rate leads to adjustment of the phonological structure or just to phonetic compression, or maybe just to a different perception by the listener. We found that listeners perceive the words as rhythmically restructured, with shifted secondary stress. Acoustically, however, we could not find any characteristic of secondary stress. What we did find was that the syllables which were perceived as carrying secondary stress were always located around 300 ms before the main stress syllable. This led us to the conclusion that listeners are equipped with an 'internal metronome'.

The second issue we discussed was the question whether recursive structures exist in phonology, just as they exist in syntax. Notably, recursion is also a well-known structural characteristic in the extralinguistic world: in nature, visual art, and music. But it would be

odd to assume that phonology does not have recursive structures, as is assumed in the Strict Layer Hypothesis (Selkirk 1984). On the basis of the repeated application of early pitch accent placement in syntactically recursive noun phrases, we showed that recursive embedding of phrases does indeed exist in phonology.

The third issue we experimentally explored concerned the question whether differences in emotional speech are characterized by different modalities. In music the difference between sad and cheerful melodies is often indicated as a difference between a minor and a major key. In a small-scale experiment with sad and cheerful stories we found indications that people also speak in a minor key when they are sad and in a major key when they are happy.

In our view, the observation that language and music show so many similarities strengthens the hypothesis that the same structures and principles hold for all temporally ordered behavior. It is the way in which our brain works: our cognitive system structures the world surrounding us in a particular way in order to understand everything in the best way.

The most striking outcome of our experiments on rhythm and recursive phrasing is that sometimes people hear things that are not present in the acoustic signal. In the case of recursive phrasing in Chapter 4, accents were perceived on the early syllable, while the only difference with the words in isolation was situated in the main stress syllable. This means that a change in one syllable may induce the perception of a change in a different syllable. This is also known as de-accenting (Horne 1990, Gussenhoven 1991). The speaker has the facilities to mark the structural entities in his speech. The listener, in his turn, can add such structural markers to his perception of the speaker's message if these markers are absent. Apparently, listeners base their perception not only on the acoustic signal alone; some strategy for retrieving the linguistic structure of the utterance also must play a role.

A similar effect of deceptive perception was found in Chapter 3 on rhythm in fast speech. Secondary stress was perceived in nearly all words, and the places of these secondary stresses were quite consistently perceived. Nevertheless, none of the acoustic measurements could indicate the secondary stresses. We found out that the listener focuses, unconsciously, on certain points in time, using an 'internal metronome', expecting information of some

importance in those positions. He therefore thinks that he hears an accent, even if it is not actually present. This again points to a communicative strategy the listener uses to anticipate the most prominent parts of the message, which he wants to extract from the sound signal.

The reason the listener uses such an anticipatory strategy is probably that human working memory is finite. One cannot focus on the whole signal altogether. Therefore, one has to extract parts of the signal, and one needs a strategy to determine which parts are essential to understanding the message. Since stress and accent are major cues of salience, the fact that listeners report that they perceive stress on a syllable they expect to be significant is therefore a logical consequence.

One can imagine that an equal tuning of the speech rate of speaker and listener would probably lead to optimal communication in both directions. The speaker will then put the most important parts of the message on the temporal locations the listener is anticipating. Regular rhythm, as Quené and Port (2005) show, leads to faster speech recognition; it enhances the perceptability of speech. Here again we find a similarity between speech and music: a regular rhythm in music is also easier to remember, and prominent chord or notes in a melody, like structure-marking accents in speech, are aligned with prominent rhythmic positions. What the studies in this dissertation show is that the idea that salience is signalled by equal timing in speech performance works the other way around too: the listener expects something of importance in the signal and therefore he actually thinks that he hears an accent.

This listener-based perspective is quite new for phonology. Of course, perception studies have always been a very important element in the discovery of phonological structure. However, the finding that an important part of this perception is based on auditory illusions, where the acoustic measurements of the syllables do not coincide with the perceived differences in prominence, is a new insight. In music theory, this perspective has been taken by several musicologists, as we indicated in Chapters 2 and 3. The conclusion that it also holds for speech rhythm shows once more that language and music share the same kinds of processes, which must partly be based on the perception by the listener. The assumption that music

and speech share some cognitive characteristics is therefore a logical one.

6.2. Future directions

In this dissertation we found that part of our communication is based on the subjective experience of the listener, not on an objective representation of the stimulus alone. This has quite some implications for the methods of research in phonology. If the interpretation of the world is influenced by the mind of the perceiver, then empirical evidence must be sought for not only acoustically, but also outside of phonetics, in psychology for instance. In other words, it is not always the case that, *'meten is weten'* as we say in Dutch, which means that 'to measure is to know' does not always apply.

This listener-based perspective, and the communicative strategies that seem to be involved in rhythm and accent perception, are a very interesting subject for future investigation. I hope I will get the opportunity to set up some experiments to look further into these strategies. The main goal would then be to identify the extra-acoustic influences on perception.