



## COMPLEMENTARY PHONOLOGY

### A THEORETICAL FRAME FOR LABELLING AN ACOUSTIC DATA BASE OF DIALOGUES

*Klaus J. Kohler*

Institut für Phonetik und digitale Sprachverarbeitung (IPDS)  
Christian-Albrechts-Universität zu Kiel  
D-24098 Kiel, Germany

#### ABSTRACT

The theoretical concept of complementarity, as known from physics, is introduced into phonology with reference to the integration of segmental and componential approaches. The principle is illustrated by examples from an acoustic data base of German dialogues.

#### I. THE CONCEPT OF COMPLEMENTARITY IN SCIENCE

Around 1927, Niels Bohr introduced the notion of complementarity into physics [1]. His aim was to take account of the fact that the concepts of classical physics were no longer able to describe the phenomena of nuclear physics adequately. Thus, depending on the experimental design, light appears as a 'wave field' on one occasion, as a 'corpuscle cloud' on another - two aspects that according to the concepts of classical physics exclude each other.

Complementarity does not just mean addition of different or even contradictory views but refers to an extension of the conceptual frame of classical physics: it is a move from an 'either-or' of the classical view to an 'as-well-as' of the quantum theoretical view. This 'as-well-as' represents a 'tertium non exclusum', an indecision between two (or possibly several) possibilities. It has the status of an objectively given state of affairs, not just of a simple lack of knowledge to resolve the indecision as was the case in classical physics. Bohr was also convinced that complementarity was a general feature of reality.

Bohr's complementarity was given a new interpretation by de Broglie [2], as a complementarity between individuality and reciprocity. It is this conceptual dyad that points far beyond physics and suggests the validity of complementarity in the macroscopic world outside physics, as already suggested by Bohr. According to this view the individuality of parts and their reciprocation when the parts are joined to a whole are complementary to each other. Systems of individualized parts may be joined to a total system by reciprocal effects between them, such as atoms to molecules, molecules to macromolecules, cells

to organisms, and in this process the individuality of parts may be lost. However, the resulting global system shows features that were not inherent properties of the partial systems. In this respect the descriptions of the partial and global systems are complementary to each other.

#### II. COMPLEMENTARITY IN PHONOLOGY

This concept of complementarity in a scientific account of the world, especially under the aspects of individuality in segmented parts and reciprocity in their serial junction can help to overcome the 'either-or' that has established itself between classical-to-generative phonology, on the one hand, and prosodic-to-autosegmental/nonlinear phonology, on the other, replacing it by an 'as-well-as', which provides greater insight into speech and language than each of the alternatives separately. There is a scientific need to pursue both views and to establish the partial indecision between them as a gain in our account of phonetics and phonology. This need becomes especially significant in dealing with the segmenting, labelling and further processing of spontaneous acoustic speech data in relation to phonological categories at the phrase level.

#### 2.1. PHONDAT: An Acoustic Data Base For German

2.1.1. Data collection In the PHONDAT and ASL/VERBMOBIL-PHONDAT projects [3,4] extensive data of spoken German in the scenarios of reading and appointment scheduling dialogues have been recorded (16 kHz, 16 bit) since 1990 and transliterated in the case of spontaneous speech [5].

2.1.2. Phonetic transcription. From the orthographic texts canonical transcriptions (lexical citation form pronunciations) were derived automatically by a TTS grapheme-phoneme conversion module [3]. The 7bit ASCII SAMPA representation is segmental phonemic (with a few modifications) [3]. These prototype text transcriptions form the basis for the application of the programme MIX [3], which allows the attribution of their canonical phonemic type symbols, one after

another, to successive non-overlapping speech signal portions in oscillographic and spectrographic presentation, with a view to segmenting and labelling them. The result is a label file that provides information about the actual pronunciation found in the speech file with reference to the canonical base, on the one hand, and to the sound wave on the other. In this process 4 ways of treating offered canonical symbols are recognised:

- (a) the symbol is kept as such: S,
- (b) the symbol is replaced by another, marked as S-S',
- (c) the symbol is deleted, marked as S-
- (d) a symbol is inserted, marked as -S.

The labelling stays basically segmental phonemic, just like the canonical prototype, with some additions, e.g. glottalization, stop releases, nasalization in the case of nasal consonant elision. But beyond these no phonetic values are marked at this stage.

2.2 Complementary Phonology in the PHONDAT Data Base Processing

2.2.1. General Aspects. The underlying phonemic framework for the description of linear phonological strings of segmental units without overlap allows the systematic and economical representation of lexical items in canonical citation form. The additional marking devices for symbol deletion, symbol change and symbol insertion (besides the default (unmarked) case of symbol preservation) make it possible to relate actually spoken words in spontaneous utterances via such speech reduction concepts as assimilation and elision to the canonical transcriptions, provided the same segmental frame is used.

The linear phonemic segmental approach is, however, not fully adequate because it cannot capture such phonological processes as 'palatalization', 'levelling of stop or

fricative strictures' etc., constituting long articulatory features, i.e. 'prosodies' [6]. It is necessary, therefore, to complement the useful but not exhaustive linear segmental model by a non-linear componential one. This inability of the linear model to deal with all relevant phonetic facts is particularly evident and frequent in instances from spontaneous speech. The following representative examples from the PHONDAT dialogue material offer complementary phonology solutions to this problem.

2.2.2 Representative dialogue examples illustrating the complementarity principle.

(a) The phrase *wahrscheinlich ein bißchen* ("probably a little", see Figure 1) provides the

actual pronunciation (IPA):  
 va:ʃaxp m+ 'bɪsˀçɛn

related to canonical SAMPA:  
 va:6#S'aInlIC QaIn+ b'IsC@n

labelled as:  
 v a:6 #S 'aI n -ma l- I- C-  
 Q- aI- n-m+ b 'I s C @- n

If the labelling were only to mark the disappearance of the phonemic segments /Iç/, the transcription would not capture the reflex of the palatal, high tongue dorsum elevation for this sequence in the preceding nasal. [ɲ] instead of [n] is thus a phonological difference resulting from contextual effects in spite of a non-realization of a string of segmental units, and since the transcription is to be phonological - not narrow phonetic in a general phonetic sense - it has to mark the difference. But it cannot do so within the segmental framework of individualized phonemic units contained in the sound system set up for German.

At this point the complementary aspect of reciprocal effects in syntagmatic concatenation

'aI n -ma l- I- C- Q- aI- n-m+

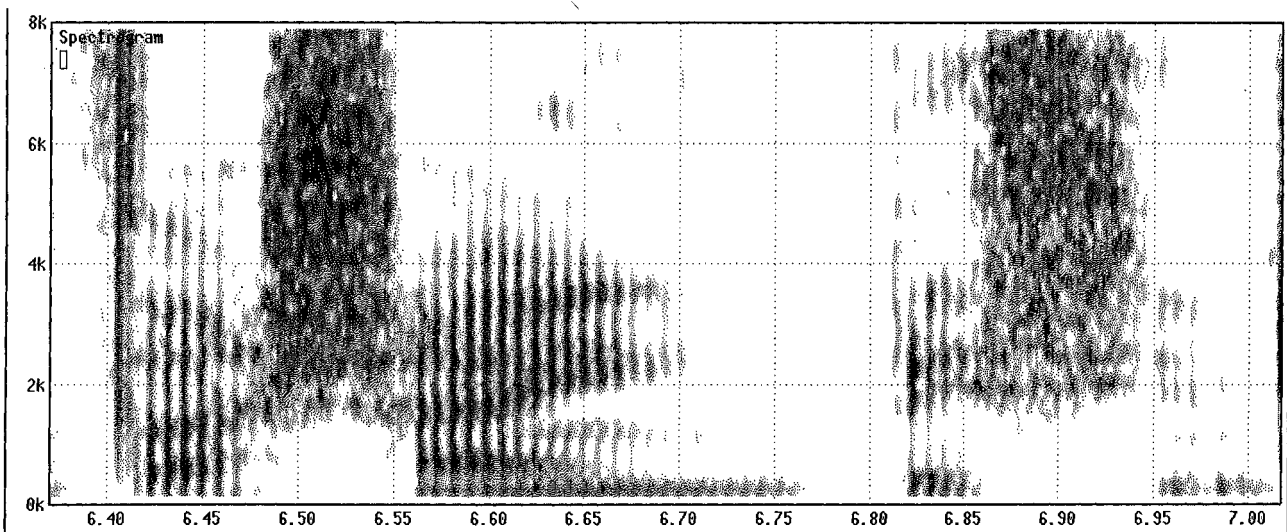
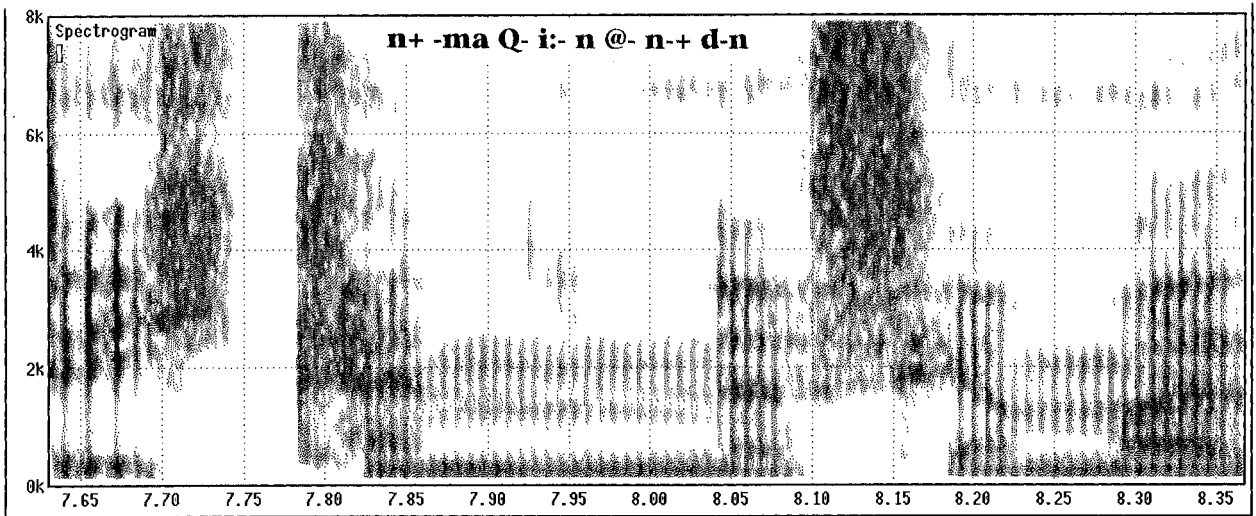
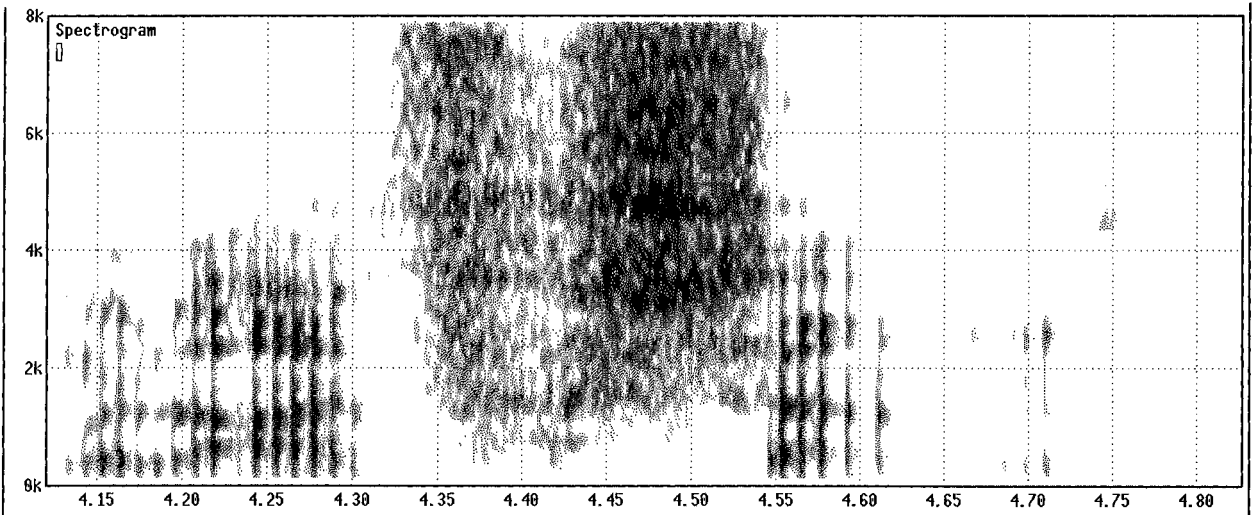


Figure 1: Spectrogram (KTH SPEG programme) of phrase (a). Labelling in SAMPA notation.



**S -ma t- "a**



**S 'aI n**

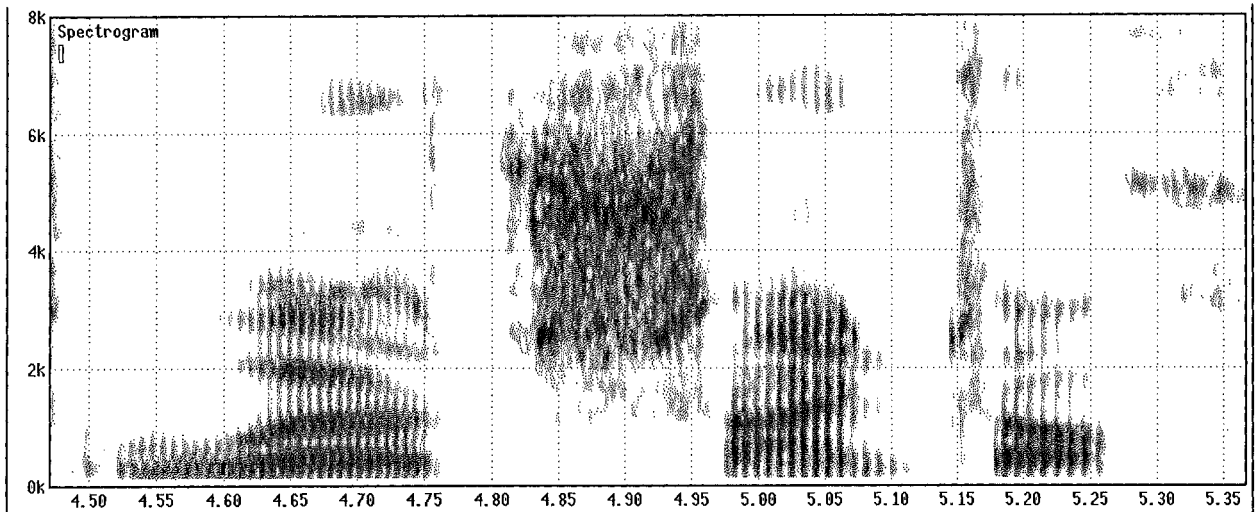


Figure 2-4: Spectrograms (KTH SPEG programme) of phrases (b), (c) and [Uhr? p:] scheint [doch]. Labelling in SAMPA notation.

tion must intervene. By postulating and symbolizing the insertion of an abstract timeless marker <ma> before the segmental units marked as deleted it is indicated that the following elided segments have left their trace - namely their feature of high dorsum elevation, i.e. palatalization or palatality - in the segment preceding the marker. The abstract marker thus has the status of a prosody without getting a particular phonetic specification itself: the complete string of 'symbol + marker + deleted symbols' (possibly interspersed with word boundary markers) determines the componential category, in this case 'palatality' combined with 'levelling of oral strictures'.

(b) In the phrase *ich kann Ihnen das ja mal [sagen]* ("I can perhaps suggest this to you") the following phonetic facts are recorded (see Figure 2):

actual pronunciation (IPA):

ɪç+ kʰan+ nʲ+ nas+ ja+ ma:l+

related to canonical SAMPA:

QIC+ kan+ Qi:n@n+das+ ja:+ ma:l+

labelled as:

Q- I C k -h a n+ -ma Q- i:- n @- n-+

d-n a s+ j a:+ m a: l+

Here the high dorsum articulation for the first vowel in *Ihnen* is transferred to the alveolar nasal consonant following it. The nasal consonant articulation is kept from the end of *kann* through *Ihnen* to the beginning of *da*, but inside it palatalization reflects the vowel of the function word *Ihnen*, although it is no longer realized as such.

(c) The phrase *einverstanden* ("agreed") shows up as (see Figure 3)

actual pronunciation (IPA):

ˈaɪnfʏʁənʦt

related to canonical SAMPA:

Q'aIn#f6St"and@n

labelled as:

Q- 'aI n#f 6 S -ma t- "a n d-Q @- n

The removed stop closure is still signalled in two ways: the fricative is not aspirated and it is without lip rounding (and palatalized). The comparison with *scheint* from the same speaker makes this clear (see Figure 4).

### III. CONCLUSIONS

The principle for the insertion of the abstract marker <ma> always is a recognizable sound difference from the simple deletions or assimilations indicated in the symbol string following the marker. The total string consisting of the marker <ma>, the symbol before it and all the symbols having the same marking device (e.g. deletion) as the first one after <ma> determines the subcategorization of the componential feature. Intervening word boundaries in the symbol string after <ma> should be considered as potential breakers of the prosodic feature. It is up to large data bank searches for data marked with the category

symbol <ma> to subcategorize excerpted strings on the strength of the left- and right-hand symbol contexts and then to replace the general marker by specific ones, representing the prosodic features involved and reflecting articulatory processes. This subcategorization will be automatic if the symbolization system and its application are systematic and unambiguous.

The complementation of the two components - segmental and prosodic - is more than their simple addition, and a representation by exclusively either the one or the other would lack a basic potential of the complementary framework: the segmental part establishes the relationship of phonetic manifestations in connected speech to canonical lexical representations, and the prosodic part introduces an overlay of articulatory processes characteristic of reduced speech. This model of complementary phonology is particularly well suited for the computational interpretation of spontaneous speech data, i.e. segmentation and labelling of acoustic signal files (mapping of symbolic and parametric descriptions), data bank searches and various computational operations on the searched strings and their corresponding signals in phonological, syntactic, semantic and pragmatic perspectives.

### ACKNOWLEDGEMENTS

Part of the work reported here was carried out with financial support from the German Federal Ministry of Research and Technology (BMFT) under ASL/VERBMOBIL contracts DLRO1IV103B4/DLRO1IV101M7. My special thanks are due to KTH/Stockholm for making MIX and related software available.

### REFERENCES

- [1] Bohr, N., "The quantum postulate and the recent development of atomic theory", *Nature* 121 (suppl.), pp. 580-590, 1928.
- [2] Broglie, de L., "Die Elementarteilchen - Individualität und Wechselwirkung". Hamburg, Goverts, 1943.
- [3] Kohler, K.J. (ed.), "Phonetisch-akustische Datenbasis des Hochdeutschen. Kieler Arbeiten zu den PHONDAT-Projekten 1989-1992". *Arbeitsberichte des Instituts für Phonetik und digitale Sprachverarbeitung (AIPUK)* 26, IPDS Kiel, 1992.
- [4] Kohler, K.J., "Lexica of the Kiel PHONDAT Corpus: Read Speech". Vols. I, II. *Arbeitsberichte des Instituts für Phonetik und digitale Sprachverarbeitung (AIPUK)* 27, 28, IPDS Kiel, 1994.
- [5] Kohler, K.J., Lex, G., Pätzold, M., Scheffers, M., Simpson, A., Thon, W., "Handbuch zur Datenaufnahme und Transliteration in TP14 von VERBMOBIL", Version 2.0. IPDS Kiel, 1994.
- [6] Firth, J.R., "Sounds and prosodies", *Transactions of the Philological Society*, pp. 127-152, 1948.