Institute of Phonetic Sciences,
University of Amsterdam,
Proceedings 23 (1999), 103-112

SUMMARIES OF PH.D. THESES DEFENDED IN 1999

TWO YEARS OF SPEECH AND LANGUAGE IN CLEFT PALATE CHILDREN (TWEE JAAR SPRAAK EN TAAL BIJ SCHISIS)

author: Kino Jansonius-Schultheiss
promotores: Louis Pols, Anne Baker
copromotor: Florien Koopmans-van Beinum
date of defence: January 19th, 1999

Summary

In this dissertation the topic of research concerns the speech and language abilities in spontaneous mother-child conversation in thirty two-year old children, born with a complete cleft or with a cleft palate only, compared with nine normal born peers. In one hand the effects of the cleft (the deviant speech motor ability caused by the cleft) on the speech and language variables was looked at, on the other hand the effect of hearing loss (the insuffficient auditory perception) and the role of medical intervention on speech and language variables was studied. Some children with a complete cleft were treated from birth on with an oral plate. Some children with a cleft and hearing loss got middle ear drains in their first year or later. Children with both forms of intervention and early treatment, in their first year of life, showed normal values on lexical and phonological tests. It is conjectured that these children with medical treatment, supporting oral tactile and auditory perception, were able to build up in time normal mental representations of speech and language, at a time when the developing brain is sensitive for sensory input, integration and cognitive linguistic organization.

It is explained in Chapter 1 that children with a complete cleft (a cleft in lip, jaw, hard and soft palate) or with a cleft palate only, run a high risk of developing speech problems. The nasal and oral passages cannot clearly be separated and this affects the articulation of both vowels and consonants. Vowels can be hypernasal and plosive consonants, which require a build-up of pressure in the mouth, are difficult to produce. Research indicates that cleft palate children are also at risk for language problems. An interesting question is what the relationship is between the development of speech at one hand, and the language

system on the other in these children and what factors influence this development. An important factor might well be hearing loss, as children with a cleft palate are at higher risk for middle ear infection (Otitis Media with Effusion or OME). Due to the cleft palate the Eustachian tube is less open and the ventilation of the middle ear is reduced.

Cleft lip and palate is obvious from birth and the children receive different types of treatment from an early age. Some babies are given an oral plate from birth on. In some children tubes are inserted in behalf of the drainage of the middle ear. All children receive surgery to close the palate (and lip and jaw) but this is done at different ages. There is no standard programme of treatment, either nationally, within the Netherlands, or internationally. It is important to know when these different treatments can best be given for optimal results in speech and language development. This development is closely related to the communicative, social, emotional, cognitive and learning development. It is therefore very important to prevent speech and language problems already at an early age.

In this study the speech and language of cleft palate children at the age of two years is examined and compared to normal development. The influence of hearing loss and medical intervention will also be evaluated.

Cleft palate can be the result of a genetic defect but this is not necessarily the case. It is relatively frequent, around two cases per thousand births (Par. 1.1). In the Netherlands the children are treated from birth, usually by a cleft palate team including an orthodontic specialist, a plastic surgeon, an ENT specialist and a speech pathologist. Some of these teams choose to insert an oral plate from birth which covers the split in the jaw and hard palate (called the Pre-surgical Orthopedic Therapy or PSOT). This treatment is applied for orthodontic purposes but from research it appears also to have beneficial effects on the production of speech sounds. The operation to close the soft palate (and sometimes at the same time the hard palate as well) is usually performed between the ages of six months and two years. Research indicates that an early operation, before seven months, is beneficial for the development of speech (Par. 1.2). The pschyosocial and cognitive development of a child can be relevant for speech and language but cleft palate children do not appear to be different in these aspects (Par. 1.3). They do show a higher prevalence of middle ear infection, however. Although hearing loss as a result of OME has not been shown to lead automatically to speech and language problems, it is possible that in combination with another problem, such as the deviant speech motor system due to the cleft palate, that it does become important (Par. 1.3). The anatomical problem does not greatly affect the production of vowels but the result is often hypernasality. Of the consonants the plosives /p/, /t/, /b/ and /d/ pose the greatest problems, the sonorants /m/, /n/, /ʋ/, /w/, /j/ and /l/ the least. The labial plosives are the most difficult to articulate, but coronal and dorsal ones are also produced less (Par. 1.4).

Chapters 2 and 3 regard both the early speech development before the first words appear (the pre-lexical stage) and the later language development. The speech and language abilities at age two years is looked upon as the result of two years of speech and language development, hearing history and medical intervention. The first year of life is important for the development of speech perception. Hearing loss as a result of OME could be crucial for later speech and language production at this time (Par. 2.1). In the first year the speech apparatus is in development (Par. 2.2). The speech production also develops quickly in this period. Speech sounds produced at this time can be considered as developmental patterns of speech movements and can be divided into movements of simple phonation, simple articulation, varied phonation, and babbling (analyzed according to Koopmans-van Beinum & Van der Stelt 1986). Varied phonation and

babbling show the greatest differences in babies with cleft palate. Babbling occurs less and with less variation as compared with peers without a cleft (Par. 2.3). The results of medical intervention in the first year are not immediately clearly observable; some results suggest a positive influence at a later age (Par. 2.4).

As a group, children with cleft palate appear to have problems with language even at school age (Par. 3.1). Therefore, it is worthwile to study speech as well as language. When the first words appear, the speech sounds are organized into a phonological system. Children with a cleft palate produce fewer sound segments correctly, particularly plosives. They are also different in phonological processes: they show more cluster reduction, more backing and more nasal assimilation (Par. 3.2). Their vocabulary appears to be smaller (Par. 3.3) and their grammatical ability is less well developed, although relatively little is known about younger children (Par. 3.4).

In Chapter 4 the methodology of this study is described. This study examines the production of speech and language in 30 cleft palate (CP) children at the age of two years. The influence of hearing loss and medical intervention is also studied (Par. 4.1). The data are taken from a spontaneous language sample of mother-child interaction (Par. 4.2). For some variables the children are compared with existing norm values; for other variables they are compared with a reference group (REF) of nine children. Both groups of children had normal intelligence (Par. 4.3.1.2). The children came from varying social-economic backgrounds (Par. 4.3.2). Both groups were comparable in medical condition except for OME and hearing (Par. 4.3.3). The evaluation of hearing loss at two years was done on the basis of field audiometry. The children were categorized in three groups of hearing loss: level 0: loss less than 20dB (CP: 2, REF: 7); level 1: loss betw. 20-40dB (CP:19, REF: 2); level 2: loss betw. 40-55dB (CP: 9, REF: 0). The children were not selected on the basis of hearing loss and are therefore not equally distributed across hearing groups. The CP children were further categorized according to whether they had received a particular medical treatment (such as an oral plate, surgical closure of the palate or a drainage of the middle ear by tubes). When a treatment had been given, it could be considered to have been applied relatively early, in the first year of the children's life, or late, thereafter. The study examines many different variables from speech production and language and also the relationship between scores on the variables (Par. 4.4). Different types of analysis are used including an analysis procedure for morphology and syntax (GRAMAT, Bol & Kuiken 1988) and for phonology (FAN, Beers 1995). Because of the type of groups non-parametric statistics are used (Par. 4.5).

Chapter 5 contains the results from general measures of speech and language, before phonology is studied in detail in Chapters 6 and 7. The CP children do not differ as a group from the REF children, neither in the amount of time they produced speech nor in the number of utterances produced, during twenty-minutes video-recording. The groups of mothers show also no differences. Communicatively the two groups are therefore comparable (Par. 5.1). The CP children can be less well understood and, not surprisingly, have more characteristics symptomatic of CP speech such as nasality (Par. 5.2). The number of spontaneous utterances and imitations the two groups produce do not differ significantly. Nor does the number of pre-lexical or lexical utterances differ (Par. 5.3). It must be noted though that the CP group shows a great variation. The cleft palate children are lexically less far developed (Par. 5.4) since as a group they produce significantly fewer content and function words (types and tokens). The analysis of grammatical development poses a methodological problem. The instrument GRAMAT requires at least 50 grammatically analyzable utterances, ideally 100. One third of the CP children cannot

be analyzed since they produce less than 50 utterances. This indicates that a subgroup exists within the CP group with a problem in language development. Of the 20 children who can be analyzed, only two show up as delayed at this age (Par. 5.5).

The analysis of the influence of hearing loss shows that it impairs intelligibility and increases CP speech as an additional factor alongside cleft palate. Hearing loss also negatively influences lexical ability and the length of the five longest utterances (MLUL) in the assessment of grammatical abilty. Children with a hearing loss between 20-40dB and who had an early placement of drainage tubes score better than children with a loss between 40-55dB or with later or no drainage (Par. 5.6). The results on all the variables show a great heterogeneity in the CP children as a group.

Chapter 6 is dedicated to the analyses of phonology on te level of correct segment production and acquired level in de complexity of contrasts within a phoneme segment. For many aspects of the analysis of phonology the FAN instrument was used (Par. 6.1). Again one third of the CP group could not be analyzed since they produced too few words; 19 of the 30 CP children could. The comparison was made with the REF children with good hearing (N=7) (Par. 6.2). As a group the 19 CP children are delayed; they achieve contrast level 2 as opposed to level 3, the level expected at the age of two years. The phonological contrast [fricative] is often absent; the contrasts [sonorant], [labial], [coronal] and [dorsal] are present. At level 2 fewer phonemes are produced than in normal children at this phonological level. Seven CP children have not acquired a single phoneme in final position which shows a large delay. Of the total of 19 CP children 10 have reached level 3 or higher but often with fewer phonemes; six children are in general delayed and three have an abnormal development. If the sound systems of the delayed and abnormal children are studied from the perspective of which sounds are almost acquired, then development can be seen to be moving for some children, but not all, in a normal direction (Par. 6.4).

Hearing has an influence on the phonological system (Par. 6.5). The children with a severe or moderate hearing loss and no early drainage had a greater chance of being excluded from the FAN analysis because they produced too few words. The CP children with a moderate loss and early drainage had the most normal phonological development. Hearing also influences the number of phonemes acquired. Hearing loss is clearly not the only factor which explains phonological delay, however. The two CP children with normal hearing both show a delayed acquisition.

Within the normal phonological acquisition the contrast grade which is reached correlates with the number of phonemes acquired. The CP score, representing the amount of deviancies in articulation due to the cleft palate, although a speech measure, does not correlate with phonological variables. Phonological acquisition is related to some other language variables. A CP child who has not acquired the contrasts [explosive] and/or [dorsal] and possesses only a limited number of phonemes, also has few types of function words, a shorter MLUL and speaks for a shorter amount of time.

In Chapter 7, also describing phonological abilities in both groups of children, the focus of research yields their phonological processes and phonotactic structure in the light of acquired hierarchy in contrast's complexity. During acquisition a child applies different types of phonological processes in word production. Segments can, for example, be deleted or substituted. Some processes are frequent in the early acquisition and are seen as normal, others are unusual and are symptomatic of a disordered acquisition. The CP group (N=19) was compared with the REF group with normal hearing (N=7) for these

analyses. The CP children show a very normal pattern in their processes, even those children with a delayed or abnormal phonology (Par. 7.5). The same processes are applied on the whole. The difference lies in more frequent use of cluster reduction and final consonant deletion (Par. 7.2), and unusual substitution processes such as backing, frication, and nasalization (Par. 7.4). The CP children often choose to produce target words which do contain the sounds which they have acquired, such as nasals. Children in the REF group however use more plosives. The cleft palate definitely has an influence on phonological processes and selection, but hearing loss is also a factor (Par. 7.6).

Chapter 8 considers the effects of medical intervention on selected speech and languge variables (Par. 8.1), in particular those which have been shown to be characteristic of CP children. The oral plate has the most influence on the phonological acquisition (Par. 8.2). Children who were not treated with an oral plate were more likely to be in the group which had such poor language that they could not be analysed. The children who were treated with the oral plate for a long time also had a better chance of having a normal contrast level and a higher number of phonemes acquired. The longer the plate had been worn, the more normal certain phonological processes were. More correct clusters were produced.

The closure of the palate was also shown to have a positive effect on intelligibility and the CP speech score. Other effects were also seen in the phonological processes (Par. 8.3). Middle ear drainage had a positive effect on the acquisition of contrasts as on certain processes. Early drainage, in the first year of baby's life, has more positive effects than later drainage (Par. 8.4). Although a factor analysis could not be executed because of the unequal distribution of the children across cells, inspection of the data strongly suggests that combined as well as early intervention has the most positive results on speech and language.

In conclusion (Chapter 9), the results from CP children at two years of age show as a group that they are statistically different in their intelligibility, the presence of CP speech deviancies and lexical ability (9.1). The heterogeneity of the group is, however, very clear. One third of the children hardly speak and are therefore hugely delayed. One third are normal in their lexical and grammatical development but have phonological problems. One third is normal in lexicon, grammar and phonology. Hearing loss due to middle ear infection has a negative effect on phonological and lexical ability. Medical intervention can alleviate the negative effects of both hearing loss and the cleft palate, especially if carried out early, in the first year of life.

The results have also produced some findings which go further than the research questions (Par. 9.2). The target sounds which the children aim at producing are different for the CP group as a whole. This finding is not interpreted as avoidance but as a result of lack of oral tactile-kinaesthetic and auditive information. When an oral plate has been placed from early on and early middle ear drainage has taken place, then the selection of these target sounds becomes more normal. It is suggested, that due to these early interventions, the development of speech perception is enhanced, leading to more normal mental representations of speech and language. The CP children are able to produce speech and language in a more normal way and in a shorter time. Another finding is that, although the mothers of the CP children as a group were not different in the amount that they spoke, the mothers of the CP children who barely speak are different. The interpretation is that this is a result of the child's slow speech and language development. This point requires further research.

The CP speech score of the children does not relate to other variables in language;

intelligibility however does relate. This suggests that the deviant phonetic production has little effect on the higher levels of organization of language (Par. 9.3).

Middle ear infection or OME is a long disputed factor in the speech and language development. This study of cleft palate children shows that it does have an effect in these children on both their phonological and lexical ability. An effect on grammatical ability is not clearly found. If children with a cleft differ in MLUL, there is an effect of the severity of hearing loss and/or the abnormal phonological development without [plosive] and/or [dorsal]. This is due to insufficient medical intervention, leading to insufficient oral speech perception and it is not necessarily due to OME. However, early drainage has a positive effect on this variable (Par. 9.4).

The study concludes with a brief look at other possible factors which might explain the behaviour in the group of CP children who barely speak at two years of age. It does not seem possible to find an explanation in the mother-child interaction, nor in the children's cognitive development. In this group of children more severe hearing loss is found (mostly a level 2 loss of 40-55 dB) with insufficient intervention. Maybe this group of children shows the effect of less language input due to more auditory deprivation than other groups of cleft palate children. The type of cleft (the cleft palate only) may be relevant. Sex of the child may also be important since most of them were boys. These are points which need further study.

PROSODIC ASPECTS OF INFORMATION STRUCTURE IN DISCOURSE

author: Monique van Donzel
promotor: Louis Pols
co-promotor: Florien Koopmans-van Beinum
date of defence: April 22nd, 1999

Summary

When producing spoken discourse, *speakers* may use various prosodic means to convey the structure of the message they are uttering. Certain words or word groups will be marked as more important than others. The whole message will furthermore be 'chunked' into smaller pieces, such as paragraphs and sentences. *Listeners* of such spoken discourse have certain ideas about the structure of the text they are hearing. Certain words or word groups will be perceived as more important than others. They are also able to detect different types of boundaries, such as sentence boundaries and paragraph boundaries. The text uttered by the speaker, the *message* itself, also has a structure. This text is assumed to be more or less coherent, otherwise the listener will have trouble understanding it, and can be divided into paragraphs, sentences, clauses, phrases, etc. Apart from the acoustic means, the speaker also has a variety of linguistic means available to indicate the structure of the message. This structure of spoken or written texts in terms of important information and boundaries can be referred to as the *information structure* of a text.

In this thesis, we focus on two main research questions. First of all, from the speaker's side, what acoustic means do speakers use to signal information structure in a spoken discourse? Secondly, from the listener's side, how are the acoustic cues, as provided by the speaker, used by the listener to detect the structure of the spoken message? In order to answer these questions, spontaneously spoken discourses were collected from speakers of standard Dutch. The discourse text was obtained by making verbatim transcriptions of the discourses. This text was then analyzed for discourse structure using a purely text-based method called 'Information Structure In Discourse' (ISID), that we developed for this purpose. The material was then viewed from the three components of the communication chain: the message, the speaker, and the listener.

The method for discourse analysis, ISID, is described in Chapter 2. On the basis of the different theories and approaches found in the literature, we developed a method suitable to analyze our spontaneous speech material. It was specifically designed to analyze *spontaneous* data. The results of a pilot experiment showed that indeed this method is useful to analyze the discourse structure of spontaneously spoken texts. Furthermore, it can be used to explain listeners' judgements on perceived prominence.

The model has a 'global' level and a 'local' level of structure. The starting point of the analysis is a verbatim transcribed spoken text, without any punctuation and/or typographical lay-out. The first step is to identify the *hierarchical level of global discourse structure* (from the largest to the smaller functional unit): discourse → paragraphs → sentences → clauses → word groups. The discourse thus consists of paragraphs, which consist of sentences, which consist of clauses, which consist of word groups. The next step is to determine the *local level of information structure*. For each word group in the clause, the information status is labeled

accordingly. The main category labels are: new, inferrable, evoked, discourse marker, modifier, and verb.

Chapter 3 gives an overview of the speakers and the material used in this thesis. The material is characterized in terms of various overall acoustic-prosodic measures, to see what means are used by the speakers to realize the essential information in spoken discourse.

Eight speakers of standard Dutch, four male and four female, were chosen. They were asked to read aloud a short story in Dutch at a normal speaking rate ('Een triomf', by Simon Carmiggelt, 1966; see appendix A and B). After a short break they were asked to retell the story they just read in their own words, with as many details as possible. The speech material was then described in temporal, pausal, and intonational features. The *temporal* aspects include clause, sentence, and paragraph durations, and speech tempo. *Pausal* features include type and duration of pauses, whereas the *intonational* aspects focus on global measures as minimum, maximum, and mean F0, range, peak heights, and high and low end tone.

A perception test was then carried out, in which listeners were asked to evaluate various prosodic aspects in the retold stories, using semantic scales. The aim was to see which features at the prosodic level listeners prefer when listening to a retold story in Dutch, and if 'good' and 'bad' speakers can be distinguished in this respect. The listeners were instructed to judge how well the speaker had performed the retelling task regarding the prosodic-acoustic realization. The results show that aspects related to the factor *Intonational characteristics* are most important, followed by the factors *Dynamic features* and *Articulation quality*. They furthermore show that there are clearly 'good' and 'bad' speakers when it comes to retelling a story. Still, all speakers were evaluated as sufficient. In general, the female speakers were preferred to the male speakers.

The next step was to see how the judgements from the perceptual evaluation test can be explained by looking at the acoustic aspects. The acoustic cue related to the judgements on *Intonational characteristics* appeared to be 'mean F0', in combination with 'minimum F0' and/or 'maximum F0'. Listeners apparently based their judgements for the various scales on the 'average F0' of the speaker, combined with either the 'maximum F0', the 'minimum F0', or both. For the *Dynamic characteristics*, all acoustic cues were relevant for the listener. A higher speaking rate, for example, caused speakers to be evaluated as more 'self-confident'. Or, the shorter a speaker realized his/her pauses, the more he/she was judged to be 'powerful'.

In conclusion, Chapter 3 shows us that the *female* speakers were evaluated as 'good' altogether. However, low correlations with the ideal speaker for temporal and pausal features lower the overall judgement score. The *male* speakers vary much more. In general, male speakers were acceptable in retelling a story, without being very appreciated.

In Chapter 4 we looked at how discourse boundaries (the global level in our ISID model) are realized acoustically by the speakers, and which of these acoustic cues are most important for the listener in the perception of discourse boundaries. A listening experiment was carried out to investigate where listeners perceive boundaries of different 'depths' in the discourse (non-final, sentence final, and paragraph final).

The data show that the relation between perceived and structural discourse boundaries is as follows: clause boundaries are, as was expected, associated with 'non-final' judgements. Sentence boundaries, however, are equally often associated with 'non-final' and with 'sentence final' judgements. Paragraph boundaries are mostly associated with 'sentence final' judgements, and not with the expected 'paragraph final' judgements. This suggests that in spoken discourse, the notion of 'paragraph' is not relevant for listeners. This seems to be a characteristic of *written* text that does not have perceptual relevance in *spoken* text, at least in the retold versions used in the present study.

The speakers realized discourse boundaries *intonationally* mainly with high boundary tones, also at places where low boundary tones were expected. Thus, in spontaneous

discourse, speakers tend to mark *continuity* at boundaries rather than *finality*. Discourse boundaries are also marked by *pausing*. Heavier and more important boundaries are marked more often with a pause, pauses are also significantly longer in duration the stronger the boundary becomes. In this sense, pausing is a more reliable cue for boundary marking than the use of boundary tones. The *combination of boundary tones and pausing*, finally, shows that the heavier the boundary, the more often both cues are used by the speaker. This confirms the hypothesis that heavier boundaries are marked by more prosodic cues than shallower ones.

Perceived boundaries were classified as either 'weak', 'strong', or 'extra strong', depending on the number and type of judgement given by the listeners. Looking then at the *acoustic characteristics of these perceived boundaries*, we saw that boundaries perceived as strong were realized with low boundary tones, whereas boundaries perceived as being less strong were realized with high boundary tones. However, pausing appeared to be more important for boundary perception than intonation. The stronger the perceived boundary, the more often it was realized with a pause, up to 100% for 'extra strong' boundaries. The *combination of pausing and boundary tones*, again, showed that the stronger a perceived boundary, the more often both prosodic cues were used by the speaker to realize this boundary.

Chapter 5 deals with the prosodic characteristics of information structure in discourse. The aim was to see how information status (the 'local' level in our ISID model) is prosodically realized by the speakers, using characteristics such as pitch accentuation and pausing. This acoustic realization of information structure was related to perceptual judgements by listeners, in this case perceived prominence.

The results for perceived prominence showed that there is a clear ordering in the relation between information status and *perceived prominence*, going from more to less prominent: new > inferrable > modifiers & verbs > discourse markers. The same ordering is found for the relation between information status and *pitch accentuation*. This same ordering is also visible if we look at the so-called *prominent pitch accents*, i.e. those words realized with a pitch accent and also perceived as being prominent. Pitch accent is the main cue for prominence perception.

However, pausing also appeared to have some relation with information status. Pauses realized *within* concepts occurred in concepts expressing new information in the discourse, i.e. new and inferrable information, and not in concepts expressing information that was already available. Pauses realized *between* concepts, i.e. on word group boundaries, did not have such a relation with information structure.

The ordering in the relation between information structure and pausing (either between or within concepts) is different from the one we saw before: discourse markers > new > inferrable > modifiers > verbs. The category *discourse markers* has moved to the first place in this ordering. This can be explained by the fact that discourse markers are located at the major turning points in the discourse. By pausing at these locations, the speaker indicates that he/she is planning the continuation of his/her story. To conclude, information structure in spoken discourse at the 'local level' is predominantly marked with intonational means, except for discourse markers, which are marked more with temporal means.

The conclusion of this thesis (Chapter 6) is that in spontaneous speech, discourse boundaries are realized with, in order of importance, silent pauses and high boundary tones. The stronger the boundary, the more often both cues are used by the speaker. Information status is realized by means of pitch accentuation, except for the category 'discourse marker', which is realized by means of pausing. Listeners make use of prosodic cues to detect information structure in discourse: the main cue for boundary perception is pausing, for information status it is the presence of a pitch accent.

In this thesis we have focused on temporal and intonational features in relation to discourse structure, from the point of view of the speaker, the listener, and the text of the discourse itself. Other prosodic aspects, such as spectral features and segmental information are important as well, but have not been included here.

Apart from the *prosodic* means, speakers did, of course, also use *linguistic* means to signal the structure of their message, such as word order and grammatical structures. An interesting follow-up of the present research would be to combine linguistic and prosodic aspects of discourse structure, and to investigate the systematic properties (pragmatic, semantic, and syntactic) of acoustic chunks uttered by the speakers of a spontaneous discourse or dialogue. How important are these properties to the listener as indicators of efficient communication? The study of dialogues between two or more speakers, or human/machine dialogues, in relation to the linguistic and prosodic aspects of these dialogues would not only give more insight into the grammatical structure of spoken discourse and dialogue, but would furthermore be of importance to language and speech technological applications such as dialogue systems.