1. Projectinformatie
Kindersprache, Aphasie und allgemeine Lautgesetze revisited

The set of contrasting sounds of a language is called the sound inventory. The Russian linguist Roman Jakobson (1941) proposed the intriguing idea that these sound inventories are structured according to a small set of universal principles, which help us understand typologies (how are sound inventories of different languages and dialects related to each other?), language acquisition (in what order do children acquire the sounds of language?) and aphasia (which sounds tend to become problematic in language impairment?). The current proposal aims to revive this program and to remedy the criticisms leveled against Jakobson, by applying insights from current phonological theory and by making use of modern database technologies and data assessing methodologies.

2. Uw gegevens
Prof. dr. Marc van Oostendorp, Meertens Instituut/KNAW and University of Leiden

3. Medeaanvragers
Dr. Clara Levelt, University of Leiden
Dr. Ineke van der Meulen, Rijndam Revalidatiecentrum, Rotterdam

4. Previous and future submissions
This program has not been submitted to other NWO programs, and there are no plans to submit it elsewhere.

5. Institutional setting
Meertens Instituut – Onderzoeksgroep Variatielinguïstiek
Universiteit Leiden – LUCL

6. Period of funding
September 1, 2008 – September 1, 2012 (4 years)

7. Composition of the Research Team
a) main applicant: Marc van Oostendorp
b) co-applicants: Clara Levelt and Ineke van der Meulen
c) researchers:
- PhD project 1: Marijn van 't Veer; currently an MA student at Leiden University (finishing his MA thesis)
- PhD project 2: *vacancy*
- Post-doc: Dr. Bert Botma; currently assistant professor at Leiden University (PhD Thesis *Phonological Aspects of Nasality* 2004; *cum laude,* and awarded AVT/Anéla Dissertation prize for best Dutch dissertation in linguistics for 2004)

d) International Advisory Board
Dr. Sergey Avrutin (University of Utrecht; language impairment)
Prof. Dr. Outi Bat-El (Tel-Aviv University; language acquisition, theoretical phonology)
Prof. Dr. Paul Boersma (University of Amsterdam; theoretical phonology; computer modeling)
Dr. Dirk-Bart den Ouden (Northwestern University; language impairment, theoretical phonology)

Prof. Dr. Gary Dell (University of Illinois at Urbana-Champaign; aphasia, phonological encoding)
Dr. Bruce Morén (University of Tromsø; theoretical phonology)
Prof. Dr. Keren Rice (University of Toronto; theoretical phonology, linguistic typology)
Dr. Mieke van de Sandt-Koenderman (Rijndam Revalidatiecentrum, Rotterdam; aphasia)
Dr. Jeroen van de Weijer (Leiden University; theoretical phonology)

8. Structure of the Proposed Research

The project consists of three subprojects:
- *Acquisition of inventories* (PhD 1, Leiden University; supervisor: Levelt/Van Oostendorp; researcher: Marijn van 't Veer)
- *Impairment of inventories* (PhD 2, Leiden University; Leiden University; supervisor: Van der Meulen/Levelt/Van Oostendorp; researcher: *vacancy*)
- *Structure of inventories* (Post-doc, Meertens Instituut; researcher: Bert Botma).

9. Description of the proposed research

Roman Jakobson's (1941) *Kindersprache, Aphasie und allgemeine Lautgesetze* (Child language, aphasia, and phonological universals) presented a revolutionary theory about the structure of underlying sound inventories of the world's languages, linking observations about language typology, language acquisition and language pathology in one uniform framework. The present project aims to renew Jakobson's original program of bringing together these different strands of research, while incorporating innovative 21st-century methodologies and technology as well as recent theoretical and empirical insights.

9.1. Jakobson's proposal: the essentials

Jakobson suggested that speech sounds are not represented as units in the mind, but consist of internal building blocks called 'distinctive features'. These cognitive entities roughly correspond to instructions to the articulatory organs or to acoustic properties of sounds; for example, the feature called 'Voice' is an instruction to let the vocal cords vibrate, and is present in voiced sounds such as *b* and *z*. Next to introducing the fundamental notion of a feature, Jakobson formulated laws of "irreversible solidarity" based on cross-linguistic observations of sound inventories. These laws specify that one phonemic contrast implies the existence of another. He formulated specific hypotheses about the order in which children acquire the segment inventory of their native language, and about the nature of language dissolution in language impairment. Children acquire the sound inventory of their native language in terms of a series of contrasts; universal, broad contrasts are acquired first: consonants versus vowels, oral versus nasal consonants, labial versus dental consonants and stops versus fricatives. Finer-grained contrasts are acquired later. According to Jakobson, the process is reversed in aphasia: fine-grained contrasts are more susceptible to loss than broader contrasts, and typologically universal contrasts are more stable than non-universal contrasts.

9.2. Some problems with Jakobson (1941)

Jakobson's program of linking phonological universals based on cross-linguistic observations to language development and language dissolution, is not as vibrant today as it started out being in 1941. Despite the great elegance and appeal of this structuralist approach, it was arguably ahead of its time: phonology as a field lacked sufficient empirical resources and theoretically refined instruments to test Jakobson's hypothesis.

The ever developing trend of specialisation, where research into theoretical phonology, linguistic typology, language acquisition and aphasia branched out into separate fields, and where researchers from one field have little access to results in the other fields, also contributed to the fading interest in a comprehensive theory.

Empirically, Jakobson's hypotheses turned out to be difficult to test (Ingram 1988), and at first sight, details of the hypotheses were not supported when they were put to the test. With respect to the order of acquisition, for example, Jakobson's predictions turned out to be too restrictive. In particular, it was found that children acquiring the same native language could have different orders of development (a.o. Braine 1971; Beers 1995). With respect to aphasia similar problems with Jakobson's hypotheses were encountered (a.o. Blumstein 1990; Béland et al. 1990; Caramazza 1994).

Theoretical developments contributed to the decline of Jakobson's approach too. With the introduction and advance of Optimality Theory (OT) in phonology (Prince & Smolensky 1993/2004) the focus shifted away from the structure of representations to constraints on outputs (Hyman 2007). A sound inventory was no longer viewed as the basic underlying set of segments of a language, but it rather emerged as the outcome of an unrestricted set of underlying segments (Richness of the Base), evaluated by ranked constraints on surface structures. As a result, the study of inventory structure has not been very prominent. Since grammars in OT evaluate individual forms, system-wide regularities tend to be neglected (but see e.g. Flemming 1995, Boersma and Hamann 2007). As such, it remains unaccounted for why sound inventories across languages do not consist of just any combination of vowels and consonants. In alternative, more phonetic or exemplar-based approaches the status of the phoneme as an abstract structural unit has also been undermined. In these approaches segments are stored in the minds of speakers/listeners in all their phonetic detail, and there is no abstract layer of representation (Bybee 2001, Pierrehumbert 2002). Sound patterns or categories of segments, if present at all, emerge diachronically from these phonetic forms (Blevins 2005). None of the current approaches therefore provide a uniform account of the structures of sound inventories in typology, acquisition and aphasia.

9.3. Towards a way out

However, recent developments make it worthwhile to revisit Jakobson's approach. First, large collections of cross-linguistic, dialectal and developmental data have become available, as have tools to search and analyze these databases. Examples of the renewed interest in databases for phonological research include the foundation of the European Corpus Phonology Group (CorPho, cofounded by the principal applicant), as well as current efforts to make accessible a large number of corpora of children's phonological development (Phon).

Secondly, promising new theoretical insights can also help solve the problems observed above. Recent work of Clements (2005, 2007) carefully evaluates new advances in phonetics and phonology in the light of a large set of segment inventories

in languages of the world, and concludes that sound inventories are, after all, best accounted for in terms of distinctive features (see also Hinskens and Van de Weijer 2003). Clements' results have been replicated in a micro-typological study of Dutch dialects (Hinskens and Van Oostendorp 2007). Regarding child language, Levelt & van Oostendorp (2007) found that the developing segment inventories of 6 young children acquiring Dutch are surprisingly well captured by an account in terms of monovalent features, underspecification, and constraints on combinations of acquired features. Phonological studies of aphasia have been sparse, but it was shown in recent work by Buchwald (2005) that the disordered production of consonant clusters by a stroke patient could be captured by the same set of OT markedness constraints that describes the developmental pattern of consonant cluster production.
Thirdly, the empirical problems noticed by Braine (1971), Beers (1995) and others, did not so much originate with the fundamental notion of a feature, but rather with the underdeveloped theory of the organization of a grammar. In particular, Levelt & van Oostendorp (2007) have shown that taking into account the notion of a feature cooccurrence constraint (FCC) – very common in theoretical phonology but neglected in research on language acquisition – many of the problems disappear.
Taken together, these developments indicate that there are promising new ways of approaching Jakobson's idea that sound inventories are structured according to well-defined principles, and that these principles apply cross-linguistically, developmentally and in aphasia.

9.4. Scientific goal and methodology of this project

The goal of this project is thus to discover the principles underlying sound inventories by means of quantitative and qualitative methods. New insights in phonology, phonetics and from production and perception studies will be taken into account. Using current refined theoretical and empirical tools, previous criticism can be countered. It is particularly important to also consider perception in the case of child language and aphasia, since production data alone cannot tell us whether problems with certain segments are due to the absence of features, or to cognitive constraints, physical limitations or limited motor control.
The three subdisciplines that are combined in this program will clearly benefit from each other. Their research traditions have become very different over the years: where experimental methodology has been scarce in phonological studies, phonological theory – as opposed to syntactic theory - has only been scantily applied to data from aphasic speech. Production data have mostly been studied separately from perception data, and studies of macro- and micro-variation have not always been successful in distinguishing one from the other. Close collaboration of the researchers involved here will therefore lead to a fruitful, and genuinely interdisciplinary approach.

Two general assumptions in this project are as follows:

A1. We assume Jakobson's model of distinctive features, albeit with a modern twist: we adopt monovalent rather than binary features. Furthermore, we are neutral at the outset of the project as to the question whether these features are universal (as Jakobson supposed) or emergent (as some more recent literature has suggested).

A2. We assume a constraint-based grammar and we will formalize our view of language variation and development in these terms. The occurrence (or non-occurrence) of segments within an inventory is therefore subject to a number of

constraints. Based on the pilot study of Levelt & van Oostendorp (2007), at present we envisage three constraint types: constraints against individual features (*F), negative feature cooccurrence constraints (*FG, i.e. a segment should not have both features F and G) and positive feature cooccurrence constraints (F>G, i.e. if a segment has F, it should also have G).

Our working hypothesis is that all sound inventories, whether from typologies, child language or aphasic speech can be described based on these assumptions about representations and constraints.
Databases in each of the three empirical areas will be tested against a set of coherent hypotheses:

H1. First, every individual inventory of natural language can be uniquely described in terms of a single set of simple constraints.

H2. Secondly, based on the above assumptions, logical relations between different types of inventories can be described straightforwardly:

H2a. In the case of acquisition, we expect that inventories of subsequent stages of development can be captured by the monotonic emergence of features and constraint rerankings.

H2b. In the case of languages of the world, differences in constraint rankings are expected to correspond to known phylogenetic and other relations between languages.

H2c. In the case of aphasia, less errors (deletions or substitutions) are expected in parts of the inventory that resemble early developmental inventories, or crosslinguistic inventories that are subsets of the inventory of the native language of the aphasic individual.

The main research questions that will be posed in each of the subprojects are:

Q1. What is the set of features needed to describe the inventories? Is there reason to believe that these features are universal?

Q2. Can every inventory be described using the minimal formalism outlined above?

Q3. In what ways do inventories differ from each other? Can these differences be described in terms of constraint rankings?

These questions will be addressed in each project by using the methodology appropriate to the project at hand (see below). By assessing and comparing the results of each of the three individual projects, in the end we hope to be able to answer these questions for the combined set of data.

9.5. Relevance and innovative qualities; academic setting

The project has several innovative qualities. In the first place, it (re-)combines insights from studies into language acquisition, aphasia, and theoretical phonology. Cross-linguistic and dialectal data from large databases are combined with detailed studies of individual and developmental sound inventories. It is innovative within each of the fields by introducing both insights and methodology from the other fields. Finally, combining production and perception data in a single study is still uncommon. A final innovative aspect of the project is that it will incorporate statistical information about distributions into grammatical theory. For this, we will invoke recent theoretical tools, such as Feature Economy (Clements 2003, 2005) and T-orders (Anttila and Andrus 2006, Prince 2006).
The tripartite structure is reflected in the main research interests of the three

applicants (theoretical phonology for Van Oostendorp, language acquisition for Levelt, aphasiology for Van der Meulen). The post-doc will be based at the Meertens Instituut, where he will be able to profit from the databases which are present there, as well as from the knowledge of database technology. The PhD students will be based at Leiden University, where they will be embedded in the graduate school of the Leiden University Center for Linguistics (LUCL) and the Leiden Institute for Brain and Cognition (LIBC).
Meertens Institute, LUCL and LIBC host a number of internationally recognized researchers who will collaborate with the present team. The advisory board consists of international experts in each of the areas involved; its members will be invited to participate in a workshop, and will also otherwise be kept up-to-date about the progress of the project.

Projects
Subproject 1. Acquisition of inventories (PhD1).
PhD1 will work on the developing sound inventory of language learners. The development of the set of contrasting sounds will be studied in production as well as in perception. In a pilot study by Levelt & van Oostendorp (2007) it was found that the set of contrasting sounds in production is initially small due to either the initial absence (or non-use) of features, or to cooccurrence constraints (FCCs) that either forbid combinations of available features or imply a particular feature in case some other feature is present. Developmental inventories can contain combinations of sounds that at first sight appear fairly random. An example comes from child 8 from the CLPF database (see below), who in production developed onset consonants in the following order: (1) p, k (2) s (3) t, l (4) m (5) f, x (6) n, _. This order can clearly not be captured in terms of a developing set of features that can be freely combined as soon as they become available. Levelt & van Oostendorp (2007) show how an account in terms of a combination of a developing feature set, underspecification, and a simple set of FCCs that become active – and can be revoked at later stages – captures the particular development of the Dutch inventory of child 8:

	Features	**Constraints**	**Predicted inventory**	**Day**
1.	[labial],[velar]		{p, k}	517
2.	[continuant]	a. *[cont, Place]	{p, k, s}	572
3.	[coronal],[lateral]	-	{p,k,s,t,l}	590
4.	[nasal]	b. [nas]>[lab]	{p,k,s,t,l,m}	608
5.	-	Revoke a.	{p,k,s,t,l,m,f,x}	636
6.	-	Revoke b.	{p,k,s,t,l,m,f,x,n,ŋ}	649

Different developmental paths taken by 5 more children acquiring Dutch are captured using the same small set of FCCs. The pilot study will be elaborated on in this subproject. The exact nature of FCCs needs to be worked out and developmental data form an important source of evidence. We need to find out how and why FCCs become active in the developing grammar and how much variation there is in active FCCs in the developmental grammars of different children. An important question arising from the pilot is what forces the variable orders of feature acquisition that became apparent in the data of the different children. This will be studied in the

longitudinal CLPF database (Levelt 1994, Fikkert 1994) containing 20,000 utterances of 12 children acquiring Dutch as their native language, re-available through CHILDES October/November 2007, complemented by longitudinal data of three more Dutch children that will be made generally available through CHILDES soon. These data will also be compared to available acquisition data from typologically unrelated languages.
Up until now the development of inventories has only been studied in production data. If the structure of an inventory results from the nature of the human language system the expectation is that the principles constraining the structure of sound inventories are not confined to production. It is therefore important to investigate whether the FCCs that play a role in production also constrain perception, and if similar developments occur. It has been proposed that the same set of markedness constraints operate both in perception and production (Boersma 1998; Tesar & Smolensky 2000; Pater 2004). Since competence often emerges in perception first, we can hypothesize that features and FCCs are active in perception before they become active in production. Therefore, infants of different age groups, starting from just before the onset of speech production will be tested in the babylab at Leiden University. The perceptual knowledge of young children can be inferred from eye-fixation measurements during the auditory presentation of a stimulus (Kemler Nelson et al. 1995). In the head turn preference procedure (HTP) a subject's eye-fixation time on a light, accompanying a loudspeaker over which an auditory stimulus is played, is measured. Significant differences in fixation or looking time that go with different auditory stimuli indicate that the dissimilarity is noticed. Jusczyk et al. (2002) developed an experimental paradigm, using the HTP, which could demonstrate the role of high ranked markedness constraints in the perception of young infants. We will apply a similar methodology to test the role of FCCs.

Subproject 2. Impairment of inventories (PhD2).
PhD2 will work on the structure of segment inventories of patients suffering from aphasia. Some studies have addressed the role of phonological markedness in aphasic speech (Blumstein 1973; Nespoulous et al. 1984; Christman 1992; see Bernhardt and Stemberger 2007 for a recent overview). However, most of the recent phonological research into aphasia has been conducted within the area of psycholinguistics. Phonological errors have been studied, plus the functional loci of these errors, but not so much the phonological representations or the phonological grammars that could give a theoretical account of these errors. Syntactic studies in aphasia have shown that with the help of theoretical tools seemingly unrelated observations can be systematically described and receive a unitary explanation (Grodzinsky 1990; Avrutin 2001; Van der Meulen 2004). In phonology, such evidence is still lacking. Noticeable exceptions are Den Ouden (2002) and Buchwald (2005), who both combine an OT model of phonology with a psycholinguistic model of speech production, and account for aspects of the phonologies of individuals with aphasia. Here we will take the same approach, further remedying the current theoretical void, by studying the linguistic principles that could govern the data, while taking into account findings from psycholinguistic and neurolinguistic studies (e.g. Dell 2004, 2007; Kurowski et al. 2007).
Like in the acquisition study, the focus will be on data from Dutch. Some data will be made available to us (the Den Ouden corpus, p.c.). However, since the amount of reliable phonological data of aphasics is not as large as it is in the other two subprojects, PhD2 will have to invest a substantial amount of time collecting relevant

data, both from production (repetition and naming experiments in addition to the collection of spontaneous speech samples) and from perception (phoneme detection studies). Patients will be recruited from the Rijndam Revalidatiecentrum, which will also make available to us a relatively large sample of recorded data. There are several types of aphasia that present with literal paraphasias – phonemic errors. Like in child language, the errors in aphasia have different sources, since different functional loci of the abstract phonological system can be damaged. We will therefore try to categorize the patients into groups with lexical retrieval deficits, phonological encoding deficits, and articulatory planning deficits. Den Ouden (2002) provides guidelines for this categorization.

Since individuals with aphasia often suffer from additional handicaps, a good measure of phoneme detection for this group is the mismatch negativity (MMN), an auditory evoked potential which can be measured in an EEG. Among others, Pettigrew et al. (2005) showed that this technique gives good results with individuals with aphasia. In collaboration with people from the EEG lab in Leiden, MMN responses will be measured in a subset of volunteering individuals with aphasia. All together, these data should provide a sufficiently detailed picture of (the stability of) phoneme inventories in aphasia.

We want to find out which parts of the Dutch sound inventory are particularly affected by aphasia, and if the relatively unaffected parts are in any way related to the developmental sound inventories of children acquiring the Dutch sound inventory, or to cross-linguistic subsets of the Dutch inventory. Crucially, theoretical tools such as FCCs are expected to play a role in these data, since they are assumed to constrain language acquisition, language pathology and language typology. We thus need to find out whether the set of FCCs constraining child language and sound inventories cross-linguistically, can account for the phonemic error patterns observed in aphasia. Buchwald (2005) demonstrated that the productions of target complex clusters of his aphasic subject were constrained by the same linguistic principles that govern the distribution of clusters cross-linguistically and in language acquisition. This is a promising result in the light of our enterprise.

The data in this project can also shed some light on the origin of FCCs. Since groups of patients with different sources of their phonemic errors participate, it might be possible to trace back specific error patterns, involving specific FCCs, to specific affected sources.

In close collaboration with PhD1 and the Postdoc, affected inventories will be fitted into the broad picture of possible inventories and their linguistic structure.

Subproject 3. Structure of inventories (Post-doc)

This project aims to bring to fruition the combination of the renewed interest in the internal structure of segments (e.g. Lombardi 2001, Morén 2003, Mielke 2004, Van Oostendorp and Van de Weijer 2005) and the success of constraint-based approaches to phonology such as Optimality Theory (Prince and Smolensky 1993/2004) by studying the theoretical implications of the notion of a feature cooccurrence constraint. This will be accomplished by uncovering a set of feature cooccurrence constraints (FCCs) that can account for linguistic macro- and microvariation in segmental inventories, integrating the insights developed in subprojects 1 and 2. Empirically, the Post-doc will concentrate on two macro-typological databases (Ian Maddieson's UPSID; and Jeff Mielke's P-Base; taking into account the problems that are known to exist for each of these) and one micro-typological database (Goeman-Taeldeman-Van Reenen Database). The goal is to design a set of features (Q1) and

FCCs (Q2) that can account for all the data in these databases, and hence has a fair chance of being able to characterize the notion of a possible segment inventory in human language, as well as the range of variation between inventories (Q3).
The methodology for this project will be as follows. First, a feature analysis needs to be found which is able to describe all the phonemes in these databases in a satisfactory way (Mielke 2004; Hall 2007; Van 't Veer, in prep.). This will lead to a full description of every individual phoneme plus a list of all necessary features. Secondly, a list of all logically possible FCCs will be created for this list of features. Thirdly, we will determine which of these constraints are necessary to describe an individual sound inventory. This will divide the original set of FCCs into three parts: those which are satisfied in all languages, those which are violated in all languages, and those which are subject to linguistic variation.
In order to interpret these results, we will turn to constraint-based approaches such as Optimality Theory or Harmonic Grammar (Pater et al. 2006). The advantage of these theories is that software tools have been developed (such as OTSoft, Praat and T-Order Generator) which are able to deal with large-scale typological comparisons. The research team has already gained some experienced with these methodologies (Van 't Veer in prep.; Hinskens and Van Oostendorp 2007). In addition, these theories are widely accepted and firmly embedded into contemporary mainstream phonological thinking.
The results from this project will aim to contribute to this mainstream as well. New insights are expected, for instance, for (i) theories of segmental markedness and (ii) the theory of possible constraints.
As to (i), a classical idea is that segmental markedness is expressed in terms of structural complexity, like the number of distinctive features a segment carries. In this view, simple segments, carrying few features, are expected to occur in many languages, and have a less restricted distribution in languages. This view is too simplistic, however. For instance, there are good reasons to assume that schwa is structurally the least marked vowels, but there are many languages without schwa (e.g. Spanish). Languages that do have schwa, usually allow this vowel in restricted contexts only (schwa only occurs in unstressed position in Dutch, Van Oostendorp 2000). The current project offers a new perspective on measuring the complexity of a segment. Rather than the number of features, the number of (relevant) FCCs that are violated by a segment might indicate complexity. Since some FCCs are implicational, they *require* features rather than prohibiting them, so that segments can be 'too simple'. It will be the task of the Post-doc to establish this measure, and test it against the databases mentioned above.
As to (ii), it is one of the well-known paradoxes of current phonological theory that it is mostly constraint-based, but that a solid theory of what constitutes a possible constraint is missing. The present study will offer a contribution here. A formally restricted set of FCCs will be established, and we will also seek to distinguish FCCs contributing to the description of natural language from irrelevant FCCs in a formal way.
The theoretical work conducted by the Post-doc will obviously also be inspired by, and inspirational to, the research performed by the two PhD students.

References

Anttila, A. and C. Andrus (2006). T-Orders. Ms, Stanford.
Avrutin, S. (2001). Linguistics and agrammatism. *GLOT International* 5, No.3
Beers, M. (1995). *The phonology of normally developing and language-impaired children*. PhD Dissertation,

University of Amsterdam, Amsterdam: IFOTT.
Béland, R., D. Caplan and J-Nespoulous (1990). The role of abstract phonological representations in word production: evidence from phonemic paraphasias. *Journal of Neurolinguistics*, 5, 125-164.
Bernhardt, B. and J. Stemberger. (2007) . Phonological impairment in children and adults. In: P. de Lacy (ed.) *The Cambridge Handbook of Phonology.* Cambridge: CUP.
Blevins, J. (2005). *Evolutionary phonology: The emergence of sound patterns*, Cambridge: Cambridge University Press.
Blumstein, S. (1973). *A phonological investigation of aphasia*. The Hague, Mouton.
Blumstein, S. (1990). Phonological deficits in aphasia: theoretical perspectives. In A. Caramazza,(ed.) *Cognitive neuropsychology and neurolinguistics: advances in models of cognitive functions and impairment*. Hillsdale: Lawrence Erlbaum Associates, 33-54.
Boersma, P. (1998). *Functional phonology: formalizing the interactions between articulatory and perceptual drives*. PhD Dissertation University of Amsterdam. The Hague: HAG.
Boersma, P, and S. Hamann (2007). The evolution of auditory contrast. Ms., University of Amsterdam and University of Utrecht. http://www.fon.hum.uva.nl/paul/papers/EvolutionOfContrast.pdf
Braine, M. (1971). The acquisition of language in infant and child. In: C. Reed (ed.) *The learning of language*. New York: Appleton-Century-Crofts, 7-95.
Buchwald, A. (2005). *Sound structure representation, repair and well-formedness: grammar in spoken language production*. PhD Dissertation, Johns Hopkins University.
Bybee, J. (2001). *Phonology and language use.* Cambridge: Cambridge University Press.
Caramazza, A. 1994. Parallels and divergences in the acquisition and dissolution of language. *Philosophical Transactions: Biological Sciences*, 346:1315, 121-127.
Christman, S. (1992). Uncovering phonological regularity in neologisms: contributions of sonority theory. *Clinical Linguistics & Phonetics* 6, 1&2, 65-76.
Clements, G. (2003). Feature economy in sound systems, *Phonology* 20.3, 287-333
- (2005). The role of features in speech sound inventories. In: Eric Raimy & Charles Cairns, eds., *Contemporary Views on Architecture and Representations in Phonological Theory*. Cambridge, MA: MIT Press, in press
- (2007). Feature Theory. Ms., Université de Paris X. http://nickclements.free.fr/featuretheory.html

Dell, G., Lawler, E., Harris, H. and J. Gordon (2004). Models of errors of omission in aphasic naming. *Cognitive Neuropsychology* 21, 2-4, 125-145.
Dell, G., Martin, N. and M. Schwartz (2007). A case series test of the interactive two step model of lexical access: Predicting word repetition from picture naming. *Journal of Memory and Language* 56, 4, 490-520.
Den Ouden, D. (2002). *Phonology in aphasia*. PhD Dissertation University of Groningen. GRODIL series no. 39.
Ferguson, C., Menn, L. and C. Stoel-Gammon (eds.) (1992). *Phonological Development. Models, Research, Implications.* Timonium: York Press.
Fikkert, P. (1994). *On the acquisition of prosodic structure*. PhD Dissertation Leiden University. HIL Dissertations in Linguistics 6, The Hague: HAG.
Flemming, E. (1995). *Auditory Representations in Phonology*. PhD Dissertation, UCLA.
Hall, D. (2007). *The role and representation of contrast*. PhD Dissertation, University of Toronto.
Hyman, L. (2007). Universals in Phonology. Ms., University of California, Berkeley.
Jakobson, R. (1941). *Kindersprache, Aphasie und allgemeine Lautgesetze.* Språkvetenskapliga Sällskapets i Uppsala Förhandlingar, Uppsala: Uppsala Universitetet. English translation: *Child Language, Aphasia and Phonological Universals.* The Hague & Paris: Mouton, 1968.
Grodzinsky, Y. (1990). *Theoretical perspectives on language deficits*. Cambridge, MIT Press.
Hinskens, F. and M. van Oostendorp. (2007). De palatalisering en velarisering van nasaal-plosief clusters in GTR. Talige, buitentalige en onderzoekerseffecten. *Taal en Tongval.*
Hinskens, F. and M. van Oostendorp. (2007). Segment inventories and T-orders. Evidence from Dutch dialects. *GLOW Newsletter*.
Hinskens, F. and J. van de Weijer. (2003). Patterns of Segmental Modification in Consonant Inventories: A Cross-Linguistic Study. *Linguistics* 41: 1041-1084.
Hyman, L. (2007). Universals in phonology. Ms., University of California, Berkeley.
Ingram, D. (1988). Jakobson revisited: some evidence from the acquisition of Polish phonology. *Lingua* 75, 55-82.
Jusczyk, P., Smolensky, P. and T. Allocco (2002). How English-learning infants respond to markedness and faithfulness constraints. *Language Acquisition* 10, 1, 31-73.
Kemler Nelson D., Jusczyk, P., Mandel, D., Myers, J., Turk, A. and L. Gerken (1995). The head-turn preference procedure for testing auditory perception. *Infant Behavior and Development* 18, 111-116.
Kurowski, K, Blumstein, S. , Palumbo, C., Waldstein, R. and M. Burton (2007). Nasal consonant production in Broca's and Wernicke's aphasics: Speech deficits and neuroanatomical correlates. *Brain and Language* 100, 3, 262-275.
Levelt, C. (1994). *On the acquisition of Place*. PhD Dissertation Leiden University. HIL dissertations in Linguistics 8. The Hague: HAG.
Levelt, C. and M. van Oostendorp (2007). Feature co-occurrence constraints in L1 acquisition. To appear in *Linguistics in the Netherlands 2007*.
Lombardi, L. (ed.). (2001). *Segmental phonology in Optimality Theory: Contraints and representations.* Cambridge: CUP.
Meulen, I. van der. (2004). *Syntactic movement and comprehension deficits in Broca's aphasia*. PhD Dissertation

Leiden University.
Mielke, J. (2004). *The emergence of distinctive features.* PhD Dissertation, Ohio State University.
Morén, B. (2003) The Parallel Structures Model of Feature Geometry. *Working Papers of the Cornell Phonetics Laboratory* 15, pp. 194-270. Ithaca, USA.
Nespoulous, J.-L., Joanette, Y., Ska, B., Caplan, D. and A. Lecours (1984). Phonological disturbances in aphasia: Is there a "markedness effect" in aphasic phonemic errors? In: F. Rose (ed.), *Advances in aphasiology*, Vol. 42: Progress in aphasiology. London, Raven Press.
Oostendorp, M. van. (2000). *Phonological projection.* Berlin: Mouton.
Oostendorp, M. van, and J. van de Weijer (eds.). (2005). *The internal structure of phonological segments.* Berlin: Mouton.
Pater, J. (2004). Bridging the gap between perception and production with minimally violable constraints. In: R. Kager, J. Pater and W. Zonneveld (eds.). *Fixing priorities: constraints in phonological acquisition*. Cambridge: Cambridge University Press.
Pater, J., C. Potts and R. Bhatt (2006). Harmonic Grammar with Linear Programming. Ms, University of Massachusetts, Amherst. ROA-872.
Pettigrew, C., Murdoch, B., Kei, J., Ponton, C., Alku, P. & H. Chenery (2005). The mismatch negativity (MMN) response to complex tones and spoken words in individuals with aphasia. *Aphasiology* 19, 2, 131-163.
Pierrehumbert, J. (2002). Word-specific phonetics. *Laboratory phonology* 7, ed. by C. Gussenhoven and N. Warner, 101-39. Berlin: Mouton de Gruyter.
Prince, A. (2006). Implication & Impossibility in Grammatical Systems: What it is & How to find it. Ms, Rutgers University.
Prince, A., and P. Smolensky. (1993/2004). *Optimality Theory: Constraint Interaction in Generative Grammar.* Rutgers University Center for Cognitive Science Technical Report **2**, 1993. Published by Blackwell, London, 2004.
Veer, M. van 't. In prep. MA Thesis, Leiden University.

Databases and software

CorPho: http://www.corpho.eu/
OTSoft: http://www.linguistics.ucla.edu/people/hayes/otsoft/
P-Base: http://aix1.uottawa.ca/~jmielke/research/pbase.html
Phon: http://childes.psy.cmu.edu/phon
Praat: http://www.fon.hum.uva.nl/praat/
T-Order Generator: http://www.stanford.edu/~anttila/research/software.html
Goeman-Taeldeman-Van Reenen Database: http://www.meertens.knaw.nl/mand/
UCLA Phonological Segment Inventory Database (UPSID): http://www.linguistics.ucla.edu/faciliti/sales/software.htm
CLPF Dutch child language database: http://childes.psy.cmu.edu/phon/ (October 2007)

10. Work program

10a. PhD1.

September 2008 - February 2009	- PhD Training LUCL/LIBC - Review of literature & available databases, etc.
March 2009 - August 2009	- Development of precise theory of features and FCCs (with PhD2 and Post-Doc) - Detailed study of CLPF database - Preparation of perception experiment - PhD Training LUCL/LIBC
September 2009 – February 2010	- First perception experiment - Detailed study of CLPF database
March 2010 – August 2010	- Writing first paper for journal publication - Assisting at preparations for workshop - Preparation of second perception experiment - Presentation at international conference
September 2010 – February 2011	- Second perception experiment - Presentation at international conference - Development of precise theory of features and

	FCCs (with PhD2 and Post-Doc)
March 2011 – August 2011	- Writing second paper for journal publication - Integration of production and perception data - Presentation at international conference
September 2011-August 2012	- Final report (writing of PhD dissertation) - Teaching course on own research

N.B. This work plan assumes that the PhD candidate works full-time; otherwise, the timing will have to be adapted accordingly. The project will then end August 2013.

10b. PhD2.

September 2008 - February 2009	- PhD Training LUCL/LIBC - Review of literature & available databases, etc. - Recruitment of patients
March 2009 - August 2009	- Development of precise theory of features and FCCs (with PhD1 and Post-Doc) - Collecting production data - PhD Training LUCL/LIBC
September 2009 – February 2010	- Collecting and interpreting production data - Preparing perception experiment 1 (EEG)
March 2010 – August 2010	- Assisting at preparations for workshop - Interpreting data from production - Writing first paper for journal publication - Perception experiment 1 - Presentation at international conference
September 2010 – February 2011	- Preparation of perception experiment 2 (EEG) - Development of precise theory of features and FCCs (with PhD2 and Post-Doc)
March 2011 – August 2011	- Perception experiment 2 - Integrating perception and production data - Presentation at international conference
September 2011-August 2012	- Writing second paper for journal publication - Final report (writing of PhD dissertation) - Teaching course on own research

N.B. We are currently applying for consent of the Medical-Ethical Committee for the experiments. This work plan assumes that the PhD candidate works full-time; otherwise, the timing will have to be adapted accordingly. The project will then end August 2013.

10c. Post-doc.

January 2009 - June 2009	- Development of precise theory of features and FCCs (with PhD1 and PhD2) - Detailed study of available macro- and micro-typological databases
July 2009 - December 2009	- Writing first paper for journal publication on FCCs

	- Developing database technology for the project (with main applicant) - Review of literature on markedness; development of theory
January 2010 – June 2010	- Writing second paper for journal publication on markedness theory - Teaching course on own research - Organisation Workshop (June 2010)
July 2010 – December 2010	- Review of literature on constraint theory - Development of precise theory of features and FCCs (with PhD1 and PhD2)
January 2011 – June 2011	- Editing papers from 2010 workshop - Writing third journal article, on constraint theory
July 2011 – December 2011	- Writing synthesizing introductory chapter

10d. General.
The core team of the project (applicants and researchers) will meet throughout the whole period at a monthly basis. PhD students will be stimulated to take a period of their training abroad, at an institute or university where specialized knowledge related to their field is available (collaborating with a member of the advisory board).

11. Word Count
(general) 1982 + (project 1) 671 + (project 2) 685 + (project 3) 702

12. Planned deliverables
The project of PhD1 will result in a PhD dissertation, of which at least two chapters will be published in *Language Acquisition* and *Journal of Child Language.* The project of PhD2 will result in a PhD dissertation, of which at least two chapters will be published in journals such as *Brain and Language* and *Journal of neurolinguistics.* The PhD theses will be supervised by all three applicants. The project of the postdoctoral researcher will result in a number of publications in journals such as *Language* and *Natural Language and Linguistic Theory.* The results of an international workshop on segmental inventories, to be held in June 2010, will be published, with a synthesizing introductory study on all aspects of this problem, coauthored by the Post-doc and the applicants.

13. Short Curriculum Vitae Principal Applicant
Marc van Oostendorp (1967) received an MA in computational linguistics from Tilburg University (1991, *cum laude,* thesis *The syntax of a language without grammar*), and defended a PhD thesis on a phonological topic at the same university (*cum laude,* dissertation *Vowel quality and phonological projection*). Since then, he has been employed at the universities of Leiden and Amsterdam, and worked as a visiting professor at many universities. Since 1999 he has been employed as a researcher at the Meertens Institute, and since 2007 he has a part-time professorship on phonological microvariation (*bijzonder hoogleraar*) at Leiden University. His work is mostly on theoretical phonology as applied to the study of language variation. The project presented here combines his interest with his ‘old’ interest in

computational models.
- 2000. *Phonological Projection*. Berlin: Mouton/de Gruyter.
- 2003. 'Comparative markedness and containment.' *Theoretical Linguistics* 29 (1/2), 65-76.
- 2005. (with B. Hermans) 'Against the sonority scale: evidence from Frankish tones.' In: Hans Broekhuis & Norbert Corver & Riny Huybregts ... [et al.] [red.] *Organizing grammar*. Mouton De Gruyter, Berlin, pp. 427-4237.
- 2005. (with J. van de Weijer) 'Phonological alphabets and the structure of the segment.' In: M. van Oostendorp & J. van de Weijer [red.] *The Internal Organization of Phonological Segments*. Mouton De Gruyter, Berlin, pp. 1-27.
- 2007. *Faithfulness in Phonology.* Ann Arbor: Equinox.

14. Summary for non-specialists (in Dutch)

Meer dan 65 jaar geleden publiceerde de Russische taalgeleerde Roman Jakobson een korte monografie onder de titel *Kindertaal, afasie en algemene klankwetten* (1941). Aan dit boek lag een intrigerend idee ten grondslag: de verzamelingen van spraakklanken (klinkers en medeklinkers) zouden in alle menselijke talen op dezelfde manier zijn geordend, en dit zou ons helpen om beter te begrijpen hoe talen op elkaar lijken en van elkaar verschillen, hoe ze door kinderen verworven worden, en wat er mis kan gaan bij afasie.

Wie een taal volledig beheerst, weet onder andere wat de klinkers en de medeklinkers van die taal zijn en hoe hij deze moet uitspreken. Die kennis is grotendeels misschien onbewust, maar wel aantoonbaar aanwezig. Sprekers van het Nederlands weten bijvoorbeeld dat hun taal de medeklinkers *p, l* en *r* hebben. Een Nederlandstalige kan bovendien het verschil tussen een l en een r onderscheiden, iets waar sprekers van bijvoorbeeld het Chinees zoals bekend moeite mee hebben. Daarenboven weten ze als ze de klinker in het Engelse *bad* horen, dat deze klank geen onderdeel uitmaakt van hun moedertaal. Deze verzameling van spraakklanken die bij een specifieke taal hoort, wordt door taalwetenschappers een 'spraakklankinventaris' genoemd.

Spraakklankinventarissen worden in minstens drie deeldisciplines van de taalwetenschap bestudeerd. In de eerste plaats zijn ze onderwerp van onderzoek voor de taaltypologie, dat wil zeggen, ze vormen een deel van het onderscheid tussen talen en dialecten, terwijl ze tegelijkertijd onderworpen zijn aan universalia – als een taal bijvoorbeeld slechts drie klinkers heeft, zullen deze normaliter de fonetische [i] (de klinker in *riep*), [a] (de klinker in (*raap*) en [u] (*roep*). In de tweede plaats, spelen spraakklankinventarissen een rol in de studie van taalverwerving. Kinderen leren de klanken van hun taal in een bepaalde volgorde. Ze beginnen met een kleine verzameling die in de loop van een paar jaar telkens wordt uitgebreid tot het volledige 'volwassen' gebied bestreken wordt. En tot slot kunnen segmentinventarissen ook bestudeerd worden vanuit het oogpunt van taalpathologie: sommige vormen van afasie kunnen ertoe leiden dat sommige patiënten klinkers en medeklinkers niet meer kunnen maken en onderscheiden.

Jakobson was de eerste die een kader voorstelde om deze drie aspecten in een uniforme theorie te begrijpen. Hij stelde voor dat taalklanken niet als onanalyseerbare eenheden ('atomen') in het hoofd vertegenwoordigd zijn, maar een interne structuur hebben (als 'moleculen'). De bouwstenen van een taalklank noemde hij 'kenmerken' (features).

Kenmerken corresponderen grofweg met instructies voor de spraakorganen. Het kenmerk 'Stem' (Voice) kan bijvoorbeeld worden geïnterpreteerd als een instructie aan de stembanden om te gaan trillen; het kenmerk maakt deel uit van klanken als *b*

en *z*. Verder liet Jakobson zien hoe taalverwerking en pathologisch taalverlies elkaar lijken te weerspiegelen, en hoe taalklanken die zeldzaam zijn in talen van de wereld, doorgaans later door kinderen verworven worden. Zodoende leren Engelstalige kinderen de klinker in *bad* nadat ze *a*, *ie* en *oe* geleerd hebben, en verliezen Engelstalige afatici (waarschijnlijk) deze uitzonderlijke klinker eerder dan de klinkers die in meer talen voorkomen.
Jakobsons voorstel wordt in de literatuur nog steeds geciteerd, maar was in een aantal welgedefinieerde opzichten zijn tijd vooruit: ten tijde van publicatie van zijn boekje ontbraken zowel voldoende empirische bronnen als theoretische verfijning om de belangrijkste hypothesen goed te toetsen. In de laatste decennia zijn deze obstakels goeddeels verdwenen, met een hernieuwde belangstelling voor fonetische meettechnieken, theoretisch-fonologische inzichten en het gebruik van grote databases voor fonologisch variatieonderzoek.
Wij menen dat de tijd gekomen is om terug te keren tot Jakobsons oorspronkelijke idee. In overeenstemming met de driedeling die in de titel van zijn boek gesuggereerd wordt, werken specialisten op het gebied van de kindertaal, de afasie en de fonologische theorie samen – drie vakgebieden die sinds de jaren veertig langzaam uit elkaar gegroeid zijn – om aan de hand van grote datasets te toetsen of alle individuele inventarissen inderdaad volgens Jakobsoniaanse richtlijnen gestructureerd zijn. Voorts zullen nieuwe gegevens worden verzameld, bijvoorbeeld om ons ervan te verzekeren dat onze bevindingen op alle drie de gebieden op een gelijke empirische basis geplaatst kunnen worden.

15. Research budget

PhD1.	
Salary	177,495
Bench fee	5,000
Subtotal	182,495
PhD2	
Salary	177,495
Bench fee	5,000
Subtotal	182,495
Post-doc (0.8)	
Salary	139,929
Bench fee	5,000
Subtotal	144,929
Replacement	0
Travel costs patients (within Netherlands)	2,300
Travel costs babies (within Netherlands)	800
Travel costs PhD's (trips abroad)	6,500
Workshop	10,000
Other (support Rijndam Revalidatiecentrum)	18,900
Incidental expenses	1,581
Subtotal	40,081
TOTAL	**550,000**