

Changing Lexical Representations in the Mental Lexicon

Paula Fikkert, KNAW/University of Nijmegen

Language is an extremely complex system, which is continuously subject to change. Despite the complexity and the vast amount of variation, children learn their mother tongue quickly and efficiently. Psychologists have offered an impressive number of functional explanations for this fact. Neurologists are discovering more about the structure of the human brain every day. Nevertheless, the structure of language remains largely invisible and can only be studied indirectly. It is ultimately the job of (psycho)linguists to discover how language is structured and how it is stored in the brain. One way of doing this is to investigate children's ability to discover the structure of language from the input which they receive and to build lexical representations in their minds in such a way that they are efficient for both the perception and the production of spoken language. Another way of studying the structure of language is to investigate the ways in which it can change. The language learner plays an important role in language change. In recent years excellent databases have become available for first language acquisition and the history of Dutch, which now provide a great opportunity to study the organisation of language in the brain through these interdisciplinary yet closely related lines of research. It is our conviction that substantial progress in our understanding of how lexical representations change can be made by combining insights from acquisition studies and historical linguistics.

Introduction

This programme addresses the classical issue of the structure of phonological representations in the lexicon from a new perspective. Its focus is on how phonological representations can undergo change (a) by looking at the way children construct and reconstruct phonological representations in the course of their development; and (b) by examining how lexical representations may change in the history of languages. Although the two types of change are of a different nature, they are closely linked as the language learner "selectively intervenes in the data, favouring those variants which best conform to the language's system" (Kiparsky 1995: 655).

To illustrate this with an example, consider the effects of Umlaut in Dutch. Umlaut is a process in which vowels are fronted in certain phonologically conditioned environments. For instance, in the Old High German word pair *gast-gesti* 'guest-guests' the plural undergoes Umlaut due to the presence of a high front vowel in the inflectional ending. After final vowel reduction Umlaut is no longer a transparent process, since the triggering factor is no longer present. German has retained Umlaut as a morphological plural marker in the nominal paradigm. In Dutch, however, nominal paradigms do not show vowel quality variation in the stem: paradigms have been levelled. An important, but neglected and unexplained fact is that the effects of Umlaut are still visible in Modern Dutch. Umlaut has led to the introduction of a marked set of vowels in Dutch: the front rounded vowels, as in *muis* /mœys/ 'mouse', *rug* /rʏx/ 'back', *muur* /myr/ 'wall' and *deur* /døʊr/ 'door'. When the source of the vowel quality variation was no longer evident for the learner, the presence of many unlauded forms led to the introduction of new phonemic distinctions in the language, the front rounded vowels. The fact that the language learners opted for a marked set of vowels, which is usually acquired much later than the set of unrounded front or rounded back vowels (Levelt 1994), indicates that markedness alone does not explain acquisition: rather the characteristics of highly valued input forms guide the child in the acquisition of lexical representations and the grammatical system. This example also shows that analogical change does not necessarily lead to simplification of lexical representations.

Lexical representations play a central role in both phonology and morphology. Of particular interest is the interface between phonology and morphology, which has recently led to two important, closely-related debates, which have involved formal phonological theory, psycholinguistics, historical phonology and acquisition studies:

- (a) storage versus computation
- (b) derivational versus constraint-based theories

In classical generative phonology, underlying representations are linked to output (surface) representations through phonological rules. Underlying representations have been argued to be necessary for two reasons: (a) to specify the inventory of contrasting phonological units, and (b) to derive variant forms of a phoneme or morpheme from one underlying base form. For instance, allophonic variation, such as the aspiration of stops in English, is never used contrastively and can be learned on the basis of distributional regularities of aspiration, taking into account the prosodic structure of words: foot-initial plosives are aspirated, as in *pet* /pɛt/ [p^hɛt]; elsewhere they are unaspirated, as *spot* /spɒt/ [spɒt].

A more complex example of variation is allomorphy. The devoicing of final voiced obstruents is part of the phonology of several languages, resulting in neutralisation of voicing distinctions in output forms with underlying final voiced obstruents. As soon as resyllabification renders the final obstruent non-final, for example through the addition of a vowel-initial plural marker, the voice distinction shows up, providing evidence for an underlying voice distinction. This is illustrated in the following examples from Dutch:

- | | | | | | | | | |
|-----|------|--------|--------|---------|--------|-----------|----------|----------|
| (1) | want | /ʋant/ | [ʋant] | 'glove' | wanten | /ʋant+ən/ | [ʋantən] | 'gloves' |
| | wand | /ʋand/ | [ʋant] | 'wall' | wanden | /ʋand+ən/ | [ʋandən] | 'walls' |

The underlying feature [voice] has therefore been assumed to be part of the underlying lexical representation. Since final devoicing in such languages is entirely predictable, there is no need to store both the singular and the plural form of lexical entries in classical generative phonology (cf. Chomsky & Halle 1968, Pinker 1999); rather, one abstract form is traditionally assumed to be stored; the alternate forms are computed in the lexicon. However, to acquire the correct underlying representation, knowledge of morphology is crucially required, since the alternation only shows up in this context, i.e. in the plural formation in (1). The literature yields very little evidence of how children deal with patterns of alternation (Macken 1995, Hayes 2000). The issue of lexical representations is in fact seldom discussed in the literature on the acquisition of phonology. Most research has focussed on children's output structures and how the child's grammar relates these to the target forms, on the assumption that the representations that children have stored in their lexicons resemble the adult target forms (e.g. Smith 1973, Levelt 1994, Gnanadesikan 1995, Pater 1999). Acquisition is thus viewed as the acquisition of the grammar of a language: how lexical representations are acquired is left unstudied. This programme aims to remedy this situation.

Storage versus computation

Pinker (1999) argues, in line with the classical generative phonology view expressed above, that information that can be computed by rules, i.e. the grammar, is not stored in the lexicon (*full-parsing* model). His main arguments come from comparing data on the acquisition of regular and irregular past tense and plural formation in English (cf. also Clahsen 1999). Others (cf. Butterworth 1983) have argued that all variant forms are listed in the lexicon (*full-listing* model); yet others have argued for a *parallel dual-route* model, in which both storage and computation is assumed (cf. Frauenfelder & Schreuder 1992, Schreuder & Baayen 1995). However, the 'storage versus computation' debate has so far largely concentrated on morphological complexity (cf. Baayen's pioneer-project at the University of Nijmegen). Phonological representations are not explicitly discussed. This project focuses on phonological representations from the perspective of language acquisition and change, rather than on the role of morphology in processing strategies in adult speakers. The issue of abstract phonological representations is not directly dependent of the choice of a full-parsing or a dual-route model: both need to make reference to the abstract phonological form of lexical items.

Whereas in classical generative phonology full-parsing was assumed, recent constraint-based theories of phonology, such as Optimality Theory, seem to implicitly accept the *full-listing* model. Arguments for this view from language acquisition and language change have been proposed. Booij (1999), for example, argues that language change provides evidence for the claim that all allomorphs, even when predictable, are stored in the lexicon. It is the task of the grammar to select the correct allomorph. He argues further that phonetic details may also be stored in the lexicon. Peperkamp & Dupoux (to appear) hold a similar view regarding language acquisition.

Contrary to this view, Lahiri & Marslen-Wilson (1991) and Lahiri & Reetz (1999) argue that the only way of dealing with variation in an efficient and effective way is by specifying lexical representations minimally: too much specification in the lexicon leads to many errors, since complete matches of input and output are seldom achieved. Moreover, less specification allows for more variation as long as input and output show no mismatches. Variation falls in between a complete match and no-mismatch. The three-way distinction between match, mismatch and no-mismatch played an important role in the account of acquisition of prosodic structure (Fikkert 1994), where it was claimed that if a child's production mismatches with the adult target form, this leads to a rapid change in the grammar (parameter settings), whereas a situation of no-mismatch (but no complete match either) could exist for a much longer period and do not lead to changes in the grammar, but often to changes in the representation of words.

In this vein, and contrasting with the full-listing hypothesis, we posit in this project the hypothesis that lexical representations are underspecified and that children's lexical representations contain even less phonological and phonetic detail than the targets: a small lexicon needs little information in order to distinguish the lexical items in perception. Acquisition involves both learning grammar and building up a (phonological) lexicon in an efficient way. This programme investigates children's cognitive abilities to detect the structure of their native language's grammar and lexicon on the basis of the input they receive. This also serves as the basis for language change, which we believe are intrinsically linked.

Derivational versus constraint-based theories

Recent developments in phonological theory have shifted the burden of explanation from restrictions on lexical representations to constraints on outputs (cf. Prince & Smolensky 1993). In Optimality Theory the phonology of a language consists of a hierarchy of universal constraints which are essentially of two types. Markedness constraints favour unmarked structures over marked ones, and Faithfulness constraints require outputs that best resemble the input structures. The interaction of the constraints determines an optimal output candidate for every possible input. If a constraint against final voiced obstruents is very highly ranked in the grammar, the output [ʋant] corresponds to both underlying representations /ʋant/ and /ʋand/. In other words, the input representation is not important. Change, during the acquisition process as well as in the history of a language, is accounted for by constraint reranking. For acquisition, it is usually assumed that Markedness constraints, which are ranked high in the initial state grammar, are demoted in the hierarchy, accounting for a gradual development towards the adult target. Lexical representations are

not in focus; they are usually assumed to be identical to those of adults (i.e. the surface forms). Acquisition entails the determination of the language-specific constraint order (cf. Gnanadesikan 1995, Tesar & Smolensky 1998).

Some proposals within Optimality Theory have even rejected the notion of underlying lexical representations and claim that constraint ranking is the only way in which knowledge of contrasts is encoded (cf. Burzio 1996, 2000, Benua 1997). In many cases the treatment of alternation is subordinated to the general characterisation of the phonological constraints in the languages. The correct English plural allomorph can be learned solely on the basis of knowledge of phonotactics, for instance. Since sequences of obstruents with different voice-qualities are not allowed in English, forms like **han[ds]* or **ca[tz]* are ungrammatical. This is not true for the kind of variation shown in (1): both *wanten* and *wanden* are grammatical. Moreover, this still does not answer the question how these input forms are represented in the lexicon: are all forms stored, or just one abstract form, with the variant forms being computed through phonological rules or constraints, or none at all?

The most important aspect of Optimality Theory has been its non-derivational nature. Crucially, /ʌnd+ən/ does not undergo resyllabification in OT, since cyclicity is banned from the theory. To account for cyclic effects in a non-derivational theory there have been two basic proposals: (a) Output-Output (O-O) correspondence (cf. McCarthy 1997; Burzio 2000), i.e. the assumption that alongside Input-Output (I-O) correspondence (cf. McCarthy & Prince 1995) there exists correspondence between related outputs; and (b) lexical phonology and morphology (LPM) with interleaving mechanisms; i.e. different domains have different co-phonologies (cf. Kiparsky 2000).

Analogical levelling

Although correspondence theory was initially introduced to deal with cyclic effects in a non-derivational grammar, it also relates to analogical or paradigmatic levelling (cf. Kiparsky 2000), since properties of one form may be transferred to another. A fundamental question is whether O-O correspondence is symmetrical or asymmetrical (the latter is named Base-Output correspondence by Kiparsky). We know too little about the direction of paradigmatic levelling and even less about the cause of the process to make hard and fast claims about this (cf. Tiersma 1982). The question which forms are used as a model for paradigmatic levelling is still open. Lahiri & Fikkert (1999) have suggested that analogical levelling tends to occur more often in inflectional paradigms than in derivational paradigms: vowel length alternations have been levelled out completely in inflectional paradigms in English, but not in derivational paradigms: *sincère* – *sincerity*.

In lexical phonology the different status of derivation and inflection is expressed by level ordering. However, there is evidence that the level ordering was of a different nature during the older stages of Germanic, where inflectional endings influenced the prosodic structure of words (cf. Lahiri & Fikkert 1999, Fikkert 2000). This is evidently not the case in the modern Germanic languages: only level I affixes influence word-prosodic structure, whereas level II affixes, among which the inflectional endings, do not. This state of affairs makes crucial reference to derivations. It is at the interface between phonology and morphology that analogical levelling and lexical restructuring are called for. The key question is what leads the language learner to reorganise the phonology-morphology interface.

Alternations in morphological paradigms have often arisen in the history of the Germanic languages as a result of Umlaut, for example. The variation in vowel quality are completely levelled out in the standard Dutch inflectional paradigms, as shown in (2), but the direction of levelling can be either way (2a vs. 2b, c), even if this requires a more marked lexical representation. However, markedness has to be evaluated in the light of the whole grammar. No levelling has taken place in certain derivational paradigms, as in (2d), *lang* ‘long’ – *lengte* ‘length’ (cf. van Loey 1954, van Bree 1977). In German the situation is different: the effects of Umlaut are still visible in the plural, where it has been morphologised (2a), and extended to words that never had an umlauting factor. This is also the situation in many Dutch dialects, particularly in the East of the Netherlands:

(2)	Gothic	Old Saxon	Middle Dutch	Mod.HG	Mod.Dutch
a.	gast - gast-i	gast - geste	gast - gaste(n)	Gast - Gäste	gast – gasten ‘guest’sg-pl
b.	fani (nom)		vēne		veen ‘peat’
c.	fanjis (gen/dat)		vēne		ven ‘pool’
d.	laggs		lanc		lang ‘long’
	laŋgiþa		leng(he)de		lengte ‘length’

This raises important questions which have never been systematically investigated, but are essential for understanding change and acquisition. How do Umlaut, gemination, open syllable lengthening and apocope interact in the different West Germanic languages and dialects? Why and how does a language learner decide to invoke paradigm levelling in certain contexts, but not in others, and what is the direction of paradigm levelling? Clearly, the interaction of phonology and morphology is involved. Moreover, many segmental processes are prosodically conditioned. With the loss of the triggering factor - the presence of a high front vowel or glide - first by vowel reduction and later by apocope, the phonological relationship between *lang* and *lengte* is no longer transparently derived from one underlying base, and this has led to lexical restructuring.

Summary

The central topics in this programme are (a) what is the nature of the lexical representations of words that human beings store in their brains? (b) how are morphologically complex words related to one another and to the base? Both issues are inextricably linked to how the child builds the lexicon and how s/he in so doing, changes it. These issues have been discussed in two prominent debates of recent years - 'storage versus computation' and 'derivation versus constraints'. However, the exact content and function of lexical representations in the language faculty has been an underresearched area. This will be investigated in two ways in this programme, in which the concept of change is central. In the first line of research, data from first language acquisition will be used to shed light on the way lexical representations are constructed and reconstructed in the course of the acquisition process, particularly at the stage at which the child is discovering allomorphic variation in the input. There are virtually no studies of this aspect of the language acquisition process, and especially not for Dutch. This research will advance our knowledge of acquisition, but more importantly, it may reveal new information on how the language faculty is organised and reorganised in the course of language development, and what characteristics of the input are highly valued and serve as a base for building up the grammatical system. These insights are of significant value for our understanding of language change, too.

The second line of research addresses more specifically the question of change in the history of Dutch. Here too, the interaction between phonology and morphology must have led to restructuring of lexical representations, given that certain kinds of variation have been completely levelled out in the language. How analogical levelling fits into formal phonological theories and particularly whether and how it can affect lexical restructuring has not been investigated, nor are there any in-depth studies relating phonological and morphological changes in the history of Dutch. This is remarkable, since internationally, this line of research has led to considerable new insight into the development of English in particular (cf. Bermúdez-Otero 1999, Lahiri & Dresher 1999, Lahiri & Fikkert 1999). Now that many data on the history of Dutch are available in electronic format, the older stages of the language can be more fruitfully and thoroughly investigated. An analysis of Dutch historical data compared with those of other Germanic languages will contribute considerably to our understanding of language change in general and of the West Germanic languages in particular. Insight into the key factors in change may in turn add a new dimension to the domain of inquiry for acquisition. The general model is thus tested against very detailed investigations. Without taking the whole grammar into consideration, we obtain only a partial picture.

Goals

The aims of the project are

- (a) detailed study of the phonological and morphological changes in various paradigms throughout the history of Dutch and the way in which paradigms did or did not undergo paradigm levelling. This will be compared with cross-linguistic results from other Germanic languages.
- (b) in-depth study of how children build up lexical representations and determine the pathways of development in the lexical representations, the phonological and morphological components of grammar, and their interaction.
- (c) to determine the nature of lexical representations in the phonology-morphology interface, the direction of change, and to identify triggers for change.
- (d) to investigate the consequences of the results for formal theories of phonology and the organisation of both the lexicon and the grammar.
- (e) to provide a fundamental theoretical basis for applied research, including clinical applications, speech technology, language teaching, etc.

This will require the active involvement of the following researchers:

- *Senior researcher* (applicant), who will use her expertise in the area of language change and acquisition to supervise the two AIO-projects, and who will embed the results of the projects in a cross-linguistic perspective by (a) comparing the history of Dutch with other West-Germanic languages, and (b) comparing the acquisition data with data from other languages. This researcher will relate the results on change and acquisition and put them in the larger perspective of on the one hand the language user's cognitive abilities to learn, produce and perceive language, and on the other formal theories of grammar. This will result in various articles and a (possibly co-authored) monograph.
- an AIO, who will investigate in detail the changes in various morphological paradigms, both derivational and inflectional, in the history of Dutch and closely related languages, particularly of West German, since their histories partly overlap. The systematic exploration of data is now possible since two important sources have in the past few years become accessible on CD-ROM: The Middle Dutch data of the Corpus Gysseling and the WNT (Woordenboek der Nederlandse Taal [Dictionary of the Dutch Language]), and possibly more corpora created at the Free University. This will result in a doctoral thesis and several articles.
- an AIO, who will investigate both the early stages of acquisition and later stages in which morphological characteristics of the language are acquired. There exists an extensive database containing longitudinal spontaneous data from 12 children acquiring Dutch (CLPF-database; Fikkert 1994, Levelt 1994). In

addition to the existing naturalistic longitudinal data (a) the collection of additional data of older children is required; (b) experimental methods will be used to supplement the spontaneous data. This will result in a doctoral thesis and several articles.

- a student-assistant to assist AIO number 2 with the data collection, the conducting of experiments and particularly the phonetic transcriptions and implementation of the transcriptions in a database (one-year appointment).

Selected references

- Benua, Laura (1997). *Transderivational identity: Phonological relations between words*. Doctoral dissertation. University of Massachusetts, Amherst.
- Bermúdez-Otero, Ricardo (1999). *Constraint interaction in language change: quantity in English and Germanic*. Doctoral dissertation. University of Manchester.
- Booij, Geert (1999). The balance between storage and computation in phonology. Ms. Free University of Amsterdam.
- Bree, Cor van (1977). *Leerboek voor de historische grammatica van het Nederlands. Klank- en vormleer met een beknopte grammatica van het Gotisch*. Groningen: Wolters Noordhof.
- Burzio, Luigi (1996). Multiple correspondence. Ms. Johns Hopkins University, Baltimore.
- Burzio, Luigi (2000). "Cycles, non-derived-environment blocking, and correspondence". In J. Dekkers et al. (eds.), *Optimality theory: Phonology, syntax, and acquisition*. Oxford: Oxford University Press. 47-87.
- Butterworth, B. (1983). Lexical representation. In B. Butterworth (ed.), *Language production (Vol II): Development, Writing and Other Language Processes*. London: Academic Press. 257-294.
- Clahsen, Harald (1999). Lexical entries and rules of language: a multidisciplinary study of German inflection. To appear in: *Behavioural and Brain Sciences* 22.
- Chomsky, Noam & Morris Halle (1968). *The sound pattern of English*. New York: Harper and Row.
- Fikkert, Paula (1994). *On the acquisition of prosodic structure*. Doctoral dissertation. University of Leiden. The Hague: Holland Academic Graphics.
- Fikkert, Paula (2000). Prosodic variation in Lutgart. In Aditi Lahiri (ed.), *Analogy, Levelling, Markedness. Principles of Change in Phonology and Morphology*. Berlin: Mouton. 301-332.
- Frauenfelder, Ulli & Robert Schreuder (1992). Constraining psycholinguistic models of morphological processing and representation: The role of productivity. In Geert Booij & Jaap van Marle (eds.), *Yearbook of Morphology 1991*. Dordrecht: Kluwer. 165-183.
- Hayes, Bruce (2000). Phonological acquisition in optimality theory: The early stages. Ms. UCLA. Submitted for a forthcoming volume on phonological acquisition and typology, edited by René Kager and Wim Zonneveld.
- Gnanadesikan, Amalia (1995). Markedness and faithfulness constraints in child phonology. Rutgers Optimality Archive 75.
- Kiparsky, Paul (1995). The phonological basis of sound change. In John Goldsmith (ed.), *The handbook of phonological theory*. Cambridge MA: Blackwell. 640-670.
- Kiparsky, Paul (2000). *Paradigmatic effects*. Cambridge MA: Cambridge University Press.
- Lahiri, Aditi & B. Elan Dresher (1999). Open syllable lengthening in West Germanic. *Language* 75: 678-719.
- Lahiri, Aditi & Paula Fikkert (1999). Trisyllabic shortening: past and present. *English Language and Linguistics* 3: 229-267.
- Lahiri, Aditi & William Marslen-Wilson (1991). The mental representation of lexical form: A phonological approach to the recognition lexicon. *Cognition* 38: 245-294.
- Lahiri, Aditi & Henning Reetz (1999). The FUL speech recognition system. 137th meeting of the Acoustical Society of America. Berlin.
- Levelt, Clara C. (1994). *On the acquisition of place*. Doctoral dissertation. University of Leiden. The Hague: Holland Academic Graphics.
- Loey, A. van (1954). *Schönfelds historische grammatica van het Nederlands. Klankleer, vormleer, woordvorming*. 5th edn. Zutphen: Thieme.
- Macken, Marlys A. (1995). Phonological acquisition. In John Goldsmith (ed.), *The handbook of phonological theory*. Cambridge MA: Blackwell. 671-696.
- McCarthy, John (1997). Sympathy and phonological opacity. Ms. University of Massachusetts, Amherst.
- McCarthy, John & Alan Prince (1999). Faithfulness and identity in prosodic morphology. In René Kager, Harry van der Hulst & Wim Zonneveld (eds.), *The prosody-morphology interface*. Cambridge: Cambridge University Press. 218-309.
- Pater, Joe (1999). From phonological typology to development of receptive and productive phonological competence: applications of minimal violation. Ms. University of Massachusetts, Amherst.
- Peperkamp, Sharon & Emmanuel Dupoux (to appear). Coping with phonological variation in early lexical acquisition. In I. Lasser (ed.), *The process of language acquisition*. Berlin: Peter Lang Verlag.
- Pinker, Steven (1999). *Words and rules. The ingredients of language*. New York: Basic Books.
- Prince, Alan & Paul Smolensky (1993). Optimality theory: Constraint interaction in generative grammar. Ms. Rutgers University.
- Schreuder, Robert & Harald Baayen (1995). Modelling morphological processing. In L. B. Feldman (ed.), *Morphological aspects of language processing*. Hillsdale, New Jersey: Lawrence Erlbaum. 131-154.
- Smith, Neil (1973). *The acquisition of phonology: A case study*. London: Cambridge University Press.
- Tesar, Bruce & Paul Smolensky (1993). Learnability in optimality theory. *Linguistic Inquiry* 29: 229-268.
- Tiersma, P. M. (1982). Local and general markedness. *Language* 58: 832-849.