Prototypes and hyperspeech:
Where are they in the grammar?

Riitta Välimaa-Blum
Université de Nice-Sophia Antipolis and CNRS-UMR 7018[*]

Abstract

If languages are fundamentally symbolic systems, as it is assumed in cognitive linguistics, then it follows that what speakers learn in the course of language acquisition are symbolic systems. It is not unequivocally established that linguistically untrained speakers even have spontaneously emerging awareness of meaningless entities like the phonemes (Liberman *et al.* 1980, Lotto and Holt 2000, Port and Leary 2005, Read *et al.* 1986, Välimaa-Blum in press). In this paper, I will argue (i) that knowledge of individual phonemes and their prototypes is metalinguistic, (ii) that the hyperspeech variants of words constitute their prototypes, and (iii) that these should be explicitly represented in the grammar. Lakoff (1993) reintroduces three levels into cognitive phonology—morpheme, word and utterance levels, which I interpret in an exemplar-theoretical framework as representing three kinds of knowledge that speakers must have of the phonology and morpholexicon of their language. The morpheme level contains an exemplar-based lexicon with all the non-automatic allomorphy and word formation principles, the word level articulates the hyperspeech forms of isolated words, and the utterance level spells out the stochastically varying hypospeech shapes of the same in continuous speech. Phonologies only having an abstract 'underlying' level and a phonetic surface have no place for the hyperspeech forms, which, however, are cognitively real to speakers. In cognitive views, the prototypes of phonemes, and hence of words as well, tend to be schematic (Langacker 1987, Mompeán-González 2004, Nathan 1996, 2006, 2007, Taylor 2003), and consequently they are never instantiated as such. I consider it unlikely that a purely abstract, non-instantiated sound shape be the prototype of a category, just as unlikely as it would be for frequency to establish them. If we accept the distinct word and utterance levels, we introduce a specific point in the grammar—the word level—that spells out the hyper-articulated, best exemplars of words.

Keywords: cognitive phonology, phoneme, prototypes, metalinguistic knowledge, hyperspeech

1   Introduction

Exemplar theory proposes that speakers memorize their experiences in terms of memory structures called exemplar clouds, which contain myriads of memorized episodes of any given category (Goldinger 1998, Johnson 1997, 2005a,b, Pierrehumbert 1994, 2001, 2002). Psycholinguistic evidence indicates that speech perception is non-analytic (Goldinger 1998), and neurolinguistic studies suggest that the phonological aspects of the mental lexicon are holistic, based on auditory percepts rather than articulation (Coleman 1998). Both types of evidence are compatible with exemplar-based approaches. In phonology, the basic exemplar unit is taken to be either the phoneme (Pierrehumbert 2001) or the word (Johnson 2005a). I have opted for the second view (Välimaa-Blum in press) and assume now that there actually are no independent memory structures of speech sounds per se but only of symbolic units.

   To illustrate the phonological composition of an exemplar-based lexical entry, we may consider the noun *demon*, the exemplar cloud of which would contain at least the following non-automatic variants: [dimən] (demon), [dimoʊn] (demoniac), and [dimɑn] (demonic), visualized in Figure 1. The subscripted ...$_n$ represents the premise that the number of memorized tokens is very large and fluctuating, for an exemplar cloud is updated every time a new episode judged to belong to it is encountered, but with time, tokens are also forgotten. It is usually assumed that items with the highest frequency form the prototypical members of a category (Johnson 1997, Nosofsky 1988, Pierrehumbert 2001, Pierrehumbert *et al.* 2000), and in Figure 1, I mark [dimən] with boldface and Italics as the presumed prototype. Exemplar clouds are taken to be granular in that, within the normal perceptual limits, those tokens deemed to be the same consolidate into one cluster-unit (Pierrehumbert 2001), and the items in Figure 1 represent this granularity.
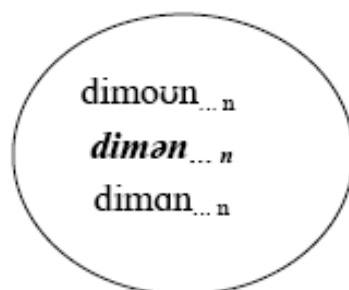

Figure 1: A granular exemplar cloud of episodes of three allomorphs of *demon*.

   I will now assume that the basic exemplar units are fully specified, meaningful sound shapes, and that the granularity in the exemplar cloud of any given lexical entry is structured in terms of its non-automatic allomorphs (Välimaa-Blum in press). There are thus no independent memory structures of speech sounds per se but phonemes are exclusively found as components of lexical entries in structures like that in Figure 1. The question that arises now is what to make out of the prototypes of phonemes that have been proposed in the literature (Jaeger 1980, Johnson 1997, Langacker 1987, Mompeán-González 2004, Nathan 2007, Nosofsky 1988, Pierrehumbert 2001, Pierrehumbert *et al.* 2000, Taylor 2003) and whether there are prototypes of phonemes at all. I will argue that any prototype and other knowledge of speech sounds as such is metalinguistic.

   Continuous speech fluctuates between hyper- and hypospeech (Lindblom 1990) and it is difficult to say which of the occurring exemplars of any given word is the typical one, and perhaps for this

reason frequency is often used as the criterion for prototypicality. I will suggest that the hyperspeech sound shapes of words are their prototypes, be they the most frequent variants or not, and also that the prototypical forms must be explicitly represented in the grammar. In the following discussion, I will use the term 'morpheme' largely in the way the term 'lexeme' is sometimes used, i.e., as an abstract unit denoting all the non-automatic allomorphs of a lexical entry, and 'word' is an actually occurring word form which is either mono- or poly-morphemic.

2 Phonemes in the representation of language

2.1 The role of phonemes and their prototypes

In the traditional generative phonology, the phoneme is an underspecified, abstract entity (Archangeli 1988a,b), and in the cognitive framework, it is taken to be schematic, i.e., also underspecified (Langacker 1987, Mompeán-González 2004, Taylor 2003), or else an 'over-specified' mental percept "fully specified for all possible features" (Nathan 2007:94). In these approaches then, the phonemes and the lexical items containing them require some kind of derivation or 'modification' when they are actually uttered. In cognitive linguistics, the schematic representation is proposed to be the prototypical member of the phoneme (Langacker 1987, Mompeán-González 2004, Taylor 2003), which means that the prototypes of phonemes are never instantiated as such.

Words and grammatical constructions all have overt instantiations and the discussion as to what the prototypical members of them are can thus be based on these actual occurrences (see, e.g., the discussion on *there*-constructions in Lakoff (1987)) and speakers can be assumed to form schematic or partly schematic knowledge of them on the basis of these episodes. If we assume now that there are prototypes of phonemes and that these are schematic abstractions that are never instantiated, it is not clear in what way they can be the prototypes. It seems problematic to attribute the status of the best or typical exemplar to an abstraction that is never observed. If prototypes were based on frequency, then also they cannot be schematic, because this would render the whole notion of frequency vacuous, for every allophone is an instantiation of the phoneme, be it schematic or not. Also, as long as we associate prototypes with representativeness or typicality, frequency is not relevant, for frequency can only establish the frequency and the degree of entrenchment of an item, not its typicality.

Whether we identify the prototypes of words with their hyper-forms or not, I maintain that prototypes must have actual instantiations for speakers to have them at all. Yet there is no principled way in the grammar to decide which of the actually occurring instantiations of any given word is its (i) hyperspeech or (ii) prototype shape. As for hyperspeech forms, it is usually just assumed that they are those variants that are fully specified and have not undergone any optional phonological processes. The fully specified hyper-forms may thus occur in speech as the result of the unification of the stem and word schema(s), but they have no special status. They are not explicitly spelled out anywhere in the grammar, for the underlying, abstract forms are underspecified or schematic, and the actual surface representation depends on its position in connected speech.

In exemplar-type memories of words, phonology is as continuous as phonetics, and therefore isolating discrete distinctive units is no more than an academic exercise, largely inspired by alphabetic writing (Port and Leary 2005, Välimaa-Blum in press). Tomasello actually argues for the "unnaturalness" of written language in general, given that humans spoke for tens of thousands of years before they started writing, and until relatively recently, the majority of human beings did not write their language at all (Tomasello 2003:3). It is also alphabetic writing that emerged last, only after the logographic and syllabic systems, which demonstrates that the analysis of speech into sounds was perhaps not the most intuitive approach in the minds of those who first started to graphically

represent language. It actually seems that even today learning to segment words into sounds is a metalinguistic achievement, not something that comes to us naturally (Liberman *et al.* 1980, Read *et al.* 1986).

Redford and Gildersleeve-Neumann (2007) suggest that in five-year-old children's clear speech, the targets are not paradigmatic but syntagmatic, i.e., children aim at reproducing adult-like *words*, not featurally distinctive sound sequences. This is clearly in agreement with the present view. Meaningfulness thus primes in that children want to produce meaning rather than that which constitutes the meaningful expression. As the procedural commands of children become fully developed, the sound sequences in their syntagmatic contexts later become more distinctive in clear speech, whereas casual speech even in adult speech remains segmentally less distinct, with the only aim at semantic clarity.

Goldinger and Azuma (2003) observe that over the last 30 years, numerous studies have tried to establish what the basic unit of speech perception is, but as of today, there is no consensus yet, and they suggest that the question of what the fundamental unit is is perhaps misguided. Their experiments indicate (i) that speech perception is simultaneously affected by both bottom-up and top-down knowledge sources and thus no single basic unit can be identified, and (ii) that episodic memory traces mediate the top-down matching (Goldinger and Azuma 2003:317). Lotto and Holt (2000) demonstrate that there is no unequivocal evidence for the role of phonemes in speech perception. They argue that the fact that experimental evidence shows that speakers are able to name the best exemplars of phonemes is not evidence of the presence of prototypical phonemes in the mental representation of language: "the presence of a "prototype" in *behavioral data* does not necessitate a "prototype" in the *mental representation* of the stimuli" [italics original—RV-B] (Lotto and Holt 2000:192).

In any case, by definition, an isolated sound contains no coarticulatory information while a contextual one does, and this would make it very difficult, already in principle, to claim that an isolated sound be the prototype of a category where the members are always contextualized. And neither can the prototype of a phoneme category be a sound occurring in some word or another, since the same phoneme is pronounced differently in different phonetic contexts. Each vowel, for example, has its own, characteristic formant (F) patterns distinguishing it from all other vowels, but these patterns change when the vowels are found in different consonantal contexts (Potter *et al.* 1966).

To see this point clearly, let us consider Figures 2 and 3. The spectrograms show an American male voice saying [æ] first in isolation and then in words. The first three formants in the isolated sound are more or less horizontal throughout, whereas in the contextual vowels, the vocal tract shapes of the surrounding consonants have modified the formant transitions differently. In all three vowels, we see F1 rise and fall at the edges. In [bæb] all the formants have largely same shape as the first, in [dæd] F2 and F3 begin with a slight fall and end as if separating, and in [gæg] the most distinctive part is the initial fall of F2 and the merging of F2 and F3 at the end. These different effects are also present in speech perception, so that if the consonants are removed from a tape recording, the listener still hears them. This is due to the fact that the vowels have integrated the consonant information in them in the form of the formant transitions. In that way, the different instances of the same vowel phoneme are dissimilar, depending on whether or not there are consonants around them, and what these consonants are.

The suggestion that the phoneme be a schematic entity overcomes this problem since it takes the phoneme to be neutered of all contextual and allophone-specific information. The schematic phoneme, however, goes against evidence indicating that speakers conserve both distinctive and non-distinctive features in the mental lexicon (Bybee 2000, 2003, Fougeron and Steriade 1997, Goldinger 1998, Johnson 1997, Miller 1994, Nosofsky 1988, Ohala and Ohala 1995, Pierrehumbert 2001, Pierre-
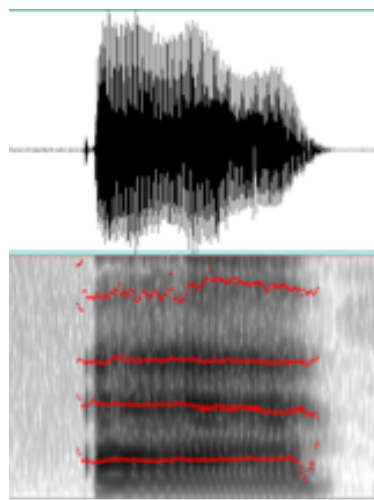
Figure 2: A spectrogram of the vowel [æ] said in isolation.
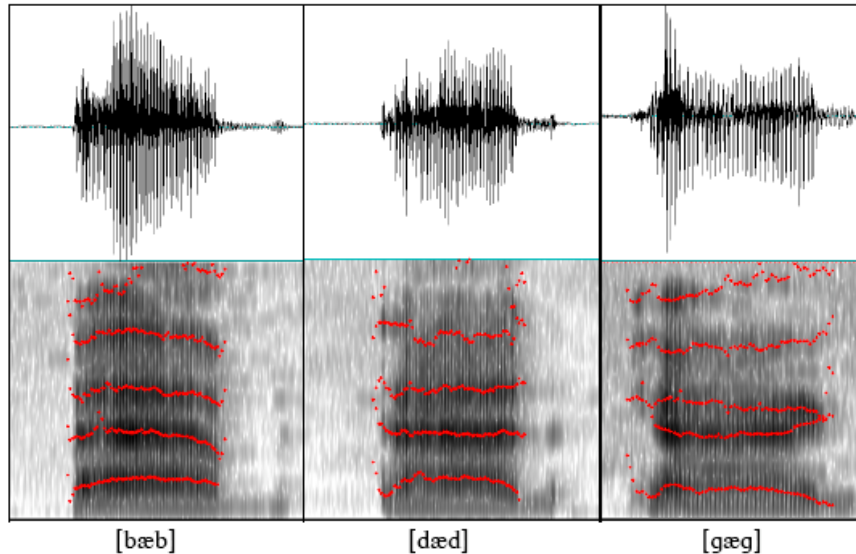

[bæb]   [dæd]   [gæg]

Figure 3: Spectrograms of the vowel [æ] in different consonant frames.

humbert *et al.* 2000) and that the lexicon that is used in both speech perception and speech production is holistic and auditory, rather than segmental and articulatory (Coleman 1998). If the lexicon thus contains forms that are fully specified for all the relevant features, it means that the lexical representation cannot be based on schematic segments. I will next argue that what speakers know of isolated phonemes belongs to metalinguistic knowledge that is explicitly learned in educational settings.

2.2 Phonemes in hyper- and hypospeech

In an experiment, Johnson *et al.* (1993) asked listeners to manipulate a speech synthesizer to produce vowels matching target vowels in words, and they found that the synthesized vowels did not match up with the models. The synthetic vowels were much more peripheral in the vowel space than the ones actually produced. The authors concluded from this that the phonetic targets of isolated

sounds correspond to hyperspeech and that actual production of the same sounds in words represents hypospeech. However, these results can also be interpreted so that vowels in words and vowels in isolation do not correspond to the same mental entities and that they do not even have identical phonetic targets. Strictly speaking, the phonetic target of a word is a word, not a sequence of sounds, and the phonetic target of an isolated sound is an isolated sound. Words are thus not made out of hypoarticulated phonemes, but both words and isolated phonemes have their own, distinct hyper- and hypo-variants.

Paradis (2002:2) discusses bilingual speakers and notes that (i) the implicit linguistic competence and metalinguistic knowledge are different in nature, and that (ii) metalinguistic knowledge is never transformed into implicit linguistic competence but remains explicitly available, and vice versa, implicit knowledge never gets converted into metalinguistic knowledge without explicit instruction. An explicitly learned language may become extremely highly entrenched and fully procedural but remains nevertheless accessible to metalinguistic, declarative statements, which is not the case of languages of which the competence is only implicit. There is no reason why this should not be true of all metalinguistic knowledge, not just of that relating to bilingual speakers. Consequently, I take the findings of Johnson *et al.* (1993) to actually concern metalinguistic knowledge of speech sounds that has been learned explicitly. The synthetic vowels generated in their experiment do not give us the hyperspeech forms of the phonemes, but only those of their metalinguistic, isolated variants.

If a language learner acquires and conserves in his long-term memory a symbolic system, then in this process, meaningless elements like the speech sounds gain no independent mental status (Välimaa-Blum in press). Symbolic units like words and grammatical constructions, however, must have an autonomous cognitive status by the very fact that they are meaningful; speakers must know their meanings and functions in order to use them in utterances that are grammatically and pragmatically well formed. On the other hand, non-symbolic units like phonemes are only constitutive parts of the symbolic elements, and there is no intrinsic motivation for speakers to have any independent knowledge of them. Speakers produce phonemes in words and utterances using highly automatized, procedural knowledge with the primary goal of expressing meaning, and knowledge of isolated sounds is only metalinguistic.

2.3   Schematic or 'ready-made' lexicon

Langacker discusses what he calls the rule/list fallacy of the generative tradition according to which grammars maximally eliminate particular statements in favor of lists of items and general rules combining them. He emphasizes that what matters in cognitive linguistics is that the complex and interconnected facts of language get adequately described, be the means general rules or specific statements (Langacker 1987:29–30). At the same time, however, Langacker proposes a model of phonology that is very much rule and list like, a view also adopted by Taylor (2003) and Mompeán-González (2004).

In this scene, speakers store schematic, underspecified phonemes in their long-term memory; these schematic phonemes are matched with schematic syllables, which in turn form schematic words (Langacker 1987:328–348). Langacker illustrates the sonority structure of syllables in terms of levels of constituency, where the vowels form an autonomous structure, and the phonologically dependent consonants are then appended to the more autonomous units (Langacker 1987:330). I of course accept this analysis as a description of the sonority patterns of syllables and words in speech, but as argued by Lotto and Holt (2000), one must not mistake "the result of organization for the cause of organization" (Koffka 1935). However, the assumption that dependent consonants are appended to more autonomous elements, and so forth for syllables and words, strangely resembles rules and lists.

While using the terminology of cognitive linguistics, the approach to cognitive phonology developed by Langacker (1987), Mompeán-González (2004), Taylor (2003) retains the essential character of generative phonology. Only now underspecified segments have become schematic, and the phonetic realization of the schematic phonemes is no longer a transformation but a modification of them. Nathan (1996, 2006), more explicitly in the framework of Natural Phonology (Donegan 1985, Donegan and Stampe 1979, Stampe 1969, 1979, 1987), takes the phoneme to be an over-specified mental percept or intention, which is phonetically realized in terms of various lenitions or fortitions. Both of these cognitive approaches, of course, also remain very close to those of, e.g., Baudouin de Courtenay (1895) and Sapir (1921), which also necessarily entail a disparity in the form between the mentally stored sound shapes and their phonetic manifestations.

If we claim, however, that languages are fundamentally symbolic systems with meaningfulness as their essential characteristic, we must rethink the role of phonology in such a system, for it cannot be the same as it is in an approach primarily focusing on linguistic form. Also, in an exemplar-based approach, there is no place for transformations or modifications, especially if lexical entries are retrieved as such from the mental lexicon (Coleman 1998). More recently, both Taylor (2006) and Mompeán-González (2006) focus less on the schematic nature of the phoneme and emphasize its taxonomic character as an assembly of its allophones, a view I have myself advocated (Välimaa-Blum 2005). Nevertheless, these authors, as well as Nathan (2006, 2007), assume phonemes to be cognitively real, basic-level categories, which, on a par with symbolic units, possess an independent mental storage base. But if the phoneme is an autonomous, basic-level category, at some level then the production of speech involves a process whereby words assembled out of schematic sound segments are transformed into fully specified segments.

If, however, we accept the evidence indicating that the mental lexicon is fully specified and auditorily based, then the need for schematic phonemes does not arise. The sounds in the fully specified lexical entries are already organized according certain principles of sonority, and they do not require any further appending or modification in the act of speaking and perceiving. They are realized according to the way the sounds are sequenced in the lexicon and the general principles that govern word formation and continuous speech in the language in question. Coleman (1998) actually suggests that the articulation of the holistic words is done "on-the-fly" using the auditorily based lexicon, and in this, schematic sound units are non-motivated, as are, of course, independent exemplar clouds of speech sounds. Speakers certainly have procedural motor-readiness for the production of speech rhythm as well as schematic knowledge of possible-word word Gestalts (Välimaa-Blum 2008), but it is unwarranted to assume that when conserving language in the long-term memory, and in speaking and listening, we in any way make use of schematic phonemes.

What undoubtedly is schematic in language are grammatical constructions (Fillmore *et al.* 1988, Goldberg 1995, 2006, Kay 1997, Lakoff 1987, Langacker 1987, Taylor 2003). Whether there are prototypical grammatical constructions in the mental representation or not raises the same issue as to whether identification equals representation (Lotto and Holt 2000), but in any case, unlike phonemes, constructions are meaningful and they must thus have an independent mental representation in their own right. If we accept Coleman's (1998) arguments that the phonological side of the mental lexicon is auditory and that words are articulatorily realized 'on-the-fly,' the schematic phonemes can be dispensed with. In actual utterances, the fully specified lexical entries are unified with meaningful, complex constructions, and depending on the speaking style, the outcome is either hyper- or hypospeech. Hyperspeech is relatively constant across contexts since it aims at the maximum of ease of perception, whereas ease of articulation dominates hypospeech, which as a result manifests a high degree of stochastic variability.

3 Where are the hyper-forms and prototypes in the grammar?

3.1 Why should hyperspeech be in the grammar?

A basic assumption in cognitive linguistics is that languages are symbolic systems, and if this is indeed so, it seems important that there be a place in the grammar that explicitly spells out the hyperspeech forms of words, which represent the smallest meaningful units, i.e., lexical entries and word schemas. These forms must have a real cognitive status in that, in addition to procedural knowledge, speakers also have propositional knowledge of the hyper-forms of words. They can repeat them at will over and over again in a relatively context-invariant manner, whereas the hypo-forms are mostly beyond conscious access. The hypo-forms manifest a high degree of stochastic variability, and on the whole, speakers only have procedural, subconscious knowledge about them. Linguistically naïve speakers may actually not even realize that there is any discrepancy between hyperspeech and hypospeech, so that when asked to repeat a word uttered in hypospeech, they typically produce the hyperspeech form.

Since declarative memories of hyperspeech sound shapes are drawn from actually observed, memorized episodes and since speakers do have conscious awareness of these forms, they should not simply be something that is only implicit in the grammar. Words, I believe, form basic-level categories and as such they should be present in the grammar, and exemplar-based approaches permit this with the fully specified, episodic lexicon. Given that spoken language in all its dialectal and stylistic variety ranges on a continuum between hyper- and hypospeech, I argue for an explicit representation in the grammar of hyperspeech forms of words, which at the same time constitute their best exemplars. In spite of their constituting basic-level categories, I have argued that words are not stored in terms of whole words, as has been suggested in the literature (Alegre and Gordon (1999), Bybee (2000), Bybee (2003:109–113), Guy (1980), Sereno and Jongman (1997), Taylor (2003:308–310)), but knowledge of word formation schemas and clusters of them is procedural (Välimaa-Blum 2005:119). In speaking then, the stems are unified 'on-the-fly' with the procedural word schemas, and the speaking style determines whether the production itself is in hyper- or hypospeech. A speaker of an unwritten language might well have less metalinguistic awareness than those who write, but even illiterate speakers certainly have some basic level of symbolic units, e.g., words, and these would certainly be more readily available in hyper- than in hypospeech.

As noted above, in the traditional generative phonology and most trends of cognitive phonology as well, there is no place where the hyperspeech forms are explicitly identified. Both approaches have an abstract, schematic or underspecified (morpho)phonemic representation and a less abstract, phonetic realization, and in this, the hyperspeech forms have no status. However, if we assume that meaningfulness is at the heart of language, it would seem important to unambiguously identify what the hyperspeech forms of the smallest units of meaning are. These units include morphologically simple word forms and partly schematic word formation schemas, which also carry meaning (Välimaa-Blum 2005). Of course, word schemas are never instantiated alone but only co-occur with their stems, and it is for this reason that both mono- and polymorphemic words stand for the lowest level of meaning, and both types of word have their own prototypes.

One motivation for representing the hyperspeech forms of words in the grammar thus comes from considerations having to do with propositional knowledge of these sound shapes and the reliable, relatively context-invariant repetitions of them. Also the fact that cognitive linguistics takes languages to be essentially symbolic systems calls for an explicit hyperspeech representation of the lowest level of symbolic function in language. Phonotactic constraints present another motivation for spelling out the hyper-forms in the grammar, for these only apply regularly to hyperspeech, not

to hypospeech. Continuous speech typically contains reductions and deletions that violate even the most basic tactic patterns. We may consider two examples from Finnish, where no word may begin or end with a consonant cluster and no native word violates this constraint in hyperspeech. In hypospeech, however, e.g., the word *mutta* [ˈmutta] 'but' can delete the entire first rhyme with its primary stress and become [mtɑ]. The question *tuletko sinä* [ˈtuletko ˈsinæ] 'are you coming?' reduces in my variety of casual hypo-Finnish to [ˈtuːks], where the two words create an illegal final cluster. An occasional loan word may contain an initial or final consonant cluster, which speakers though typically simplify, but no native word can begin with a nasal and a stop, and no Finnish word ends with a stop and a fricative. These typical instances from continuous hypospeech form a clear argument for explicitly representing hyperarticulated, isolated words in the grammar, for otherwise it is difficult to state any viable phonotactic constraints at all.

3.2   How many levels?

Lakoff (1993) reintroduces three levels into cognitive phonology, (i) the morpheme, (ii) phoneme or word and (iii) phonetic or utterance levels, without any derivation between the three, but only various constraints. In an exemplar-based framework, we can interpret these three levels as representing three different kinds of knowledge as follows (cf. Välimaa-Blum (2005:30)):

> ***Morpheme level:*** Exemplar-based mental lexicon containing all the non-automatic contextual variants of morphemes, and the word formation schemas.
>
> ***Phonemic or word level:*** Morphemes alone or unified with word schemas form words, and speakers have explicit, propositional knowledge of their fully specified hyperspeech shapes.
>
> ***Phonetic or utterance level:*** Morphemes alone or unified with word schemas form words, which occur in utterances manifesting a great deal of stochastic variability; knowledge of this variability is largely subconscious, procedural.

In the view I espouse, the morpheme level contains the mental lexicon with each lexical entry being an exemplar cloud of its non-automatic allomorphs, as illustrated above. Here we find all the phonological shapes of morphemes that the speaker has encountered and categorized, as well as the fully or partly specified words schemas, which are meaningful as well. Both the word and utterance levels spell out actually observed sound shapes of the fully specified and auditorily memorized lexical targets and word schemas. The word level represents propositional knowledge that speakers have of each word form in terms of their hyper-articulation. The more procedural utterance level then represents sound shapes that, vis-à-vis the word level, have undergone optional deletions, reductions and sandhi in continuous speech. This means that the word level describes knowledge of the phonological word shapes that are relatively invariant, including syllabification and stress, whereas the utterance level manifests a large amount of subconscious stochastic variability, thus potentially being at variance with respect to the word level instantiations, and not only in terms of segmental content but stress and syllabification as well.

I assume that syllabification per se is not present in the lexicon, but the exemplar-type lexical entries themselves are of course organized according to specific sonority patterns, alternating between high and low points, which creates the side effect of a syllabified lexicon. In Levelt's model of speech production, "syllabification is a late process in phonological encoding (in particular there is no syllabification in the word form lexicon) and…gestural scores for syllables are retrieved as whole entities" (Levelt and Wheeldon 1994:25). Syllabification is thus a late, on-line process, not present in

the mental lexicon, and this is consistent with the assumption that the lexicon is holistic, containing exemplar memories of the auditorily stored sound shapes, and that articulation is created 'on-the-fly.' The actual syllabification of isolated hyperspeech words faithfully replicates the sonority patterns of the lexical entries unified with word schemas, but the syllabification of connected speech within and across word boundaries manifests a large amount of variability. I thus consider syllabification proper to be present only in actual instantiations of words (word level) and utterances (utterance level), not in the mental lexicon.

The fact that I evoke three levels as Lakoff does, does not mean that there is any kind of derivation or 'modification,' but the levels are totally void of sequentiality, rather they are more like the mental lexicon with two types of instantiations, the hyper- and hypo-forms. The morpheme and word levels can be related to one another by regular phonological principles, but the relation between the word and utterance levels is more complex. To a certain extent, they can also be related to one another through regular phonological principles, but not in all instances, for the observed variability in continuous speech cannot always be accounted for in phonological terms, and not all is universal either. There certainly are universal articulatory principles that govern connected speech processes (Kohler 1990, Nolan 1996), but as noted by Iivonen (1996) and Butcher (1996), there are also reduction principles that seem to be language-specific, depending on the typological characteristics of the language in question. Semantic and pragmatic principles play a role as well (Kohler 1990, Välimaa-Blum 1998). Given the large amount of variability at the utterance level, it is desirable that the hyperspeech forms be explicitly present in the grammar, for they are relatively invariant and they are the forms that speakers have conscious knowledge of, whereas both the mental lexicon and hypospeech are beyond explicit awareness.

3.3   Prototypes of words coincide with their hyperspeech forms

Assuming now that we explicitly represent the hyperspeech words in the grammar, as outlined above, the next question concerns the prototypes and their representation in the grammar. If we accept that the best exemplars of words correspond to their actual instantiations, the prototype then is either the hyper-form or one of the hypospeech variants. Lakoff (1987:84–90) and Croft and Cruse (2004:80–81) discuss ways of conceptualizing the prototype members of categories in general, and among them, they mention typicality or representativeness. In phonology, apart from the schematic presentations, the prototypes of speech sounds are based on frequency (Johnson 1997, Nosofsky 1988, Pierrehumbert 2001, Pierrehumbert *et al.* 2000). However, if prototypes of speech sounds belong to metalinguistic knowledge, as I argued above, then only the prototypes of words are concerned at this point. If frequency were the central criterion for prototypicality, the most frequent token of any given word might well be one of its hypospeech variants. For example, a speaker may say *send* [sɛn] more frequently than [sɛnd], but would this be sufficient for the former to be the typical representative of the category?

Since the variability in hypospeech is stochastic rather than predictable, it would seem unlikely for any hypo-form to constitute the best exemplar. The hyperspeech forms represent very careful articulations where the articulatory targets are maximally reached, whereas hypospeech represents the case where these are reached to varying degrees, and this explains why there is necessarily more variability in hypo- than in hyperspeech. It is difficult to say whether or not hyperspeech is more frequent than hypospeech, but whichever the case, when asked to repeat a word, speakers seem to master the hyper-variants with greater consistency than the hypo-forms.

Frequency is an essential element in the entrenchment of language, but we have to dissociate it from representativeness, for there is no reason to suppose that the most frequent form is, in the

minds of the speakers, the typical member of the category. I propose that each word form has its own prototype which is the hyperspeech articulation, be it the most frequent instance or not. Function words belong to grammatical constructions rather than the lexicon, but we may assume their prototype pronunciations in a language like English to be the so-called full forms since these are the variants that speakers can readily repeat. Globally, this means that each word form has its own best exemplar, but if we further desire that each lexical entry have its prototype as well, this probably is the non-derived, uninflected form.

So, if we assume that the prototypes must have actual episodes in spoken language and if we want to explicitly represent the prototypes of words in the grammar, the decision is between the hyper-form and any one of the hypo-pronunciations. Hyperspeech variants are absent in the traditional generative and cognitive phonologies, which only assume (i) an 'underlying,' schematic level and (ii) a surface level of actually occurring words. In these types of grammar, there is no systematic way to spell out hyperspeech, which can only be identified extraneously. If we accept that Lakoff's levels represent three kinds of lexically-phonologically relevant knowledge, then we are equipped with a way to describe (i) both the hyperspeech forms of words, which I assume at the same time establish their best exemplars, and (ii) the stochastic variability found in connected utterances. Both the word and utterance levels represent actual instantiations of words and in that sense, there can be no directionality implied between them, but both are more concrete than the morpheme level, which is purely abstract.

4   Concluding remarks

In naturally occurring language use, phonemes are always and only found in meaningful units, and in the mental representation of language, there is no motivation for them to constitute independent memory structures. The prototypes and hyper-articulated forms of isolated phonemes belong to metalinguistic knowledge learned in formal settings, often explicitly related to orthography, which is another metalinguistic achievement. As for words, they all have their hyperarticulated variants, which speakers can repeat at will, and which as such represent explicit, propositional knowledge. Morphologically simple and complex words belong to the lowest level of meaningfulness in language given that they involve meaningful roots and meaningful word formation principles, and I claim that their prototypes must be based on actual episodes for speakers to learn them at all. If the prototypes of words were schematic, it would be difficult to attribute representativeness to them, for is not clear in what sense they would be typical of their categories if they are never instantiated as such. Also, if prototypes were based on frequency, then obviously the schematic forms are out of the question, and also, frequency might well result for the prototype being a hypoarticulated form. Since the hyperarticulated forms embody not only procedural but propositional knowledge as well, I suggest that they constitute the prototypes. I thus explicitly assume that prototypes in phonology be interpreted as the best or typical representatives, not as the most frequent instances.

In cognitive linguistics with its emphasis on meaningfulness in language, it is important to explicitly identify in the grammar the best exemplars of all semantic entities, including words. To do this, I adopted and adapted Lakoff's traditional three-level approach, where the word level gives the relatively invariant hyper-forms of isolated words and the utterance level their stochastically varying sound shapes as they occur in continuous speech. What is of central importance now is the primacy of meaningfulness in language and the fact that the meaningless units are always only subsidiary. Explicit knowledge of phonemes does have its place in educated speakers' linguistic knowledge base together with knowledge of orthography, but it does not belong to the implicit linguistic compe-

tence. Globally then, the mental lexicon is fully specified, based on auditory percepts, and in the act of speaking, one or another of the non-automatic allomorphs of lexical entries is retrieved as such to co-occur with word formation schemas and other grammatical constructions. The duality of structure in language separates the meaningless sounds from the meaningful elements, and this is explicit in a grammar that fully subordinates sounds to meaning, without giving them any independent existence.

References

Alegre, Maria and Peter Gordon (1999). Frequency effects and the representation status of regular inflections. *Journal of Memory and Language* 40: 41–61.

Archangeli, Diana B. (1988a). Aspects of underspecification theory. *Phonology* 5(2): 183–208.

Archangeli, Diana B. (1988b). *Underspecification in Yawelmani phonology and morphology*. New York: Garland Publishers.

Butcher, Andrew (1996). Some connected speech phenomena in Australian languages: universals and idiosyncrasies. In Adrian P. Simpson and Matthias Pätzold (eds.), *Sound patterns of connected speech: Description, models and explanation*, AIPUK 31, Kiel: Arbeitsberichte, pp. 83–104.

Bybee, Joan L. (2000). The phonology in the lexicon: Evidence from lexical diffusion. In Michael Barlow and Suzanne Kemmer (eds.), *Proceedings of the Rice Symposium on Usage-Based Models of Language*, Stanford, CA: CSLI Publications, pp. 65–85.

Bybee, Joan L. (2003). *Phonology and language use*. Cambridge: Cambridge University Press.

Coleman, John (1998). Cognitive reality and the phonological lexicon: A review. *Journal of Neurolinguistics* 11(3): 295–320.

Baudouin de Courtenay, Jan (1895). *Versuch einer Theorie phonetischer Alternationen: ein Capitel aus der Psychophonetik*. Strasburg: KJ Trübner.

Croft, William and D. Alan Cruse (2004). *Cognitive linguistics*. Cambridge: Cambridge University Press.

Donegan, Patricia J. (1985). *On the Natural Phonology of vowels*. Garland Publishers.

Donegan, Patricia J. and David Stampe (1979). The study of Natural Phonology. In Daniel A. Dinnsen (ed.), *Current approaches to phonological theory*, Bloomington: Indiana University Press, pp. 126–173.

Fillmore, Charles J., Paul Kay and Mary O'Connor (1988). Regularity and idiomaticity in grammatical constructions: The case of 'let alone'. *Language* 64(3): 501–538.

Fougeron, Cecile and Donca Steriade (1997). Does deletion of French schwa lead to neutralization of lexical distinctions? *EUROSPEECH* : 943–946.

Goldberg, Adele E. (1995). *Constructions: A Construction Grammar approach to argument structures*. University of Chicago Press.

Goldberg, Adele E. (2006). *Constructions at work: The nature of generalization in languages*. Oxford: Oxford University Press.

Goldinger, Stephen D. (1998). Echoes of echoes? an episodic theory of lexical access. *Psychological Review* 105: 251–279.

Goldinger, Stephen D. and Tamiko Azuma (2003). Puzzle-solving science: the quixotic quest for units in speech perception. *Journal of Phonetics* 31(3-4): 305–320.

Guy, Gregory (1980). Variation in the group and the individual: The case of final stop deletion. In William Labov (ed.), *Locating Language in Time and Place*, New York: Academic Press, pp. 1–36.

Iivonen, Antti (1996). Reductions in Finnish related to lexicon, grammar, stress and speech rates. In Adrian P. Simpson and Matthias Pätzold (eds.), *Sound patterns of connected speech: Description, models and explanation*, AIPUK 31, Kiel: Arbeitsberichte, pp. 105–112.

Jaeger, Jeri (1980). *Categorization in phonology: An experimental approach*. Ph.D. thesis, University of California, Berkeley.

Johnson, Keith (1997). Speech perception without speaker normalization. In Keith Johnson and John Mullennix (eds.), *Talker variability in speech processing*, San Diego, CA: Academic Press, pp. 145–165.

Johnson, Keith (2005a). Decisions and mechanisms in exemplar-based phonology. *UC Berkeley Phonology Lab Annual Report* : 289–311.

Johnson, Keith (2005b). Resonance in an exemplar-based lexicon: The emergence of social identity and phonology. *UC Berkeley Phonology Lab Annual Report* : 95–128.

Johnson, Keith, Edward Flemming and Richard Wright (1993). The hyperspace effect: Phonetic targets are hyperarticulated. *Language* 69(3): 505–28.

Kay, Paul (1997). *Words and the grammar of context*. Stanford, CA: CSLI Publications.

Koffka, Kurt (1935). *Principles of gestalt psychology*. New York: Harcourt Brace.

Kohler, Klaus (1990). Segmental reduction in connected speech in German: Phonological facts and phonetic explanations. In William J. Hardcastle and Alain Marchal (eds.), *Speech production and speech modeling*, Dordrecht: Kluwer Academic Publishers, pp. 69–92.

Lakoff, George (1987). *Women, fire, and dangerous things: what categories reveal about the mind*. Chicago: University of Chicago Press.

Lakoff, George (1993). Cognitive phonology. In John Goldsmith (ed.), *The last phonological rule: Reflections on constraints and derivations*, Chicago: University of Chicago Press, pp. 117–145.

Langacker, Ronald W. (1987). *Foundations of Cognitive Grammar: Theoretical prerequisites*, vol. 1. Stanford, CA: Stanford University Press.

Levelt, Willem J. M. and Linda Wheeldon (1994). Do speakers have access to a mental syllabary? *Cognition* 50: 239–269.

Liberman, Isabelle Y., Alvin M. Liberman, Ignatius Mattingly and Donald P. Schankweiler (1980). Orthography and the beginning reader. In James F. Kavanagh and Richard L. Venezky (eds.), *Orthography, reading, and dyslexia*, Baltimore: University Park Press, pp. 137–153.

Lindblom, Björn (1990). Explaining phonetic variation: A sketch of the H&H theory. In William J. Hardcastle and Alain Marchal (eds.), *Speech production and speech modeling*, Amsterdam: Kluwer Academic Publishers, pp. 403–439.

Lotto, Andrew J. and Lori L. Holt (2000). The illusion of the phoneme. In Sabrina J. Billings, John P. Boyle and Aaron M. Griffith (eds.), *The Panels*, Chicago: Chicago Linguistic Society, CLS vol. 35, pp. 191–204.

Miller, Joanne (1994). On the internal structure of phonetic categories: a progress report. *Cognition* 50: 271–285.

Mompeán-González, José A. (2004). Category overlap and neutralization: The importance of speakers' classifications in phonology. *Cognitive Linguistics* 15(4): 429–469.

Mompeán-González, José A. (2006). The phoneme as a basic-level category. *International Journal of English Studies* 6(2): 141–172.

Nathan, Geoffrey S. (1996). Steps towards a Cognitive Phonology. In Bernhard Hurch and Richard A. Rhodes (eds.), *Natural Phonology: the state of the art*, Berlin: Mouton de Gruyter, pp. 107–120.

Nathan, Geoffrey S. (2006). Is the phoneme usage-based?—some issues. *International Journal of English Studies* 6(2): 173–195.

Nathan, Geoffrey S. (2007). Where is the Natural Phonology phoneme in 2007? *ICPhS XVI* : 93–98.

Nolan, Francis (1996). Overview of English connected speech processes (CSPs). In Adrian P. Simpson and Matthias Pätzold (eds.), *Sound patterns of connected speech: Description, models and explanation*, AIPUK 31, Kiel: Arbeitsberichte, pp. 15–26.

Nosofsky, Robert M. (1988). Similarity, frequency, and category representation. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 14(1): 54–65.

Ohala, John J. and Manjari Ohala (1995). Speech perception and lexical representation: the role of vowel nasalization in Hindi and English. In Bruce Connell and Amalia Arvaniti (eds.), *Phonology and phonetic evidence*, Cambridge: Cambridge University Press, Papers in Laboratory Phonology vol. IV, pp. 41–60.

Paradis, Michel (2002). Neurolinguistics of bilingualism and the teachings of languages. In *P. Bouissac (Chair), The multimodality of human communication: Theories, problems, and applications. Symposium conducted at the University of Toronto, Victoria College, Canada.*

Pierrehumbert, Janet B. (1994). Knowledge of variation. In Katharine Beals, Jeannette Denton and Robert Knippen (eds.), *Variation and Linguistic Theory, 30th Meeting of the Chicago Linguistic Society*, Chicago: Chicago Linguistic Society, Chicago Linguistic Society vol. 2, pp. 232–256.

Pierrehumbert, Janet B. (2001). Exemplar dynamics: word frequency, lenition and contrast. In Joan Bybee and Paul Hopper (eds.), *Frequency and the emergence of linguistic structure*, Amsterdam: John Benjamins, pp. 137–157.

Pierrehumbert, Janet B. (2002). Word-specific phonetics. In Carlos Gussenhoven and Natasha Warner (eds.), *Laboratory phonology 7*, Berlin: Mouton de Gruyter, pp. 101–139.

Pierrehumbert, Janet B., Mary E. Beckman and D. Robert Ladd (2000). Conceptual foundations of phonology as a laboratory science. In Noel Burton-Roberts, Philip Carr, and Gerard Docherty (eds.), *Phonological knowledge: Conceptual and empirical issues*, Oxford: Oxford University Press, pp. 273–303.

Port, Robert F. and Adam P. Leary (2005). Against formal phonology. *Language* 81(4): 927–964.

Potter, Ralph K., George A. Kopp and Harriet Green Kopp (1966). *Visible speech.* New York: Dover Publications.

Read, Charles, Yun-Fei Zhang, Hong-Yin Nie and Bao-Qing Ding (1986). The ability to manipulate speech sounds depends on knowing alphabetic writing. *Cognition* 24: 31–44.

Redford, Melissa A. and Christine E. Gildersleeve-Neumann (2007). Children's clear speech suggests word-level targets: Preliminary evidence. In *ICPhS XVI*, Saarbrücken, pp. 365–368.

Sapir, Edward (1921). *Language: An introduction to the study of speech.* New York: Harcourt Brace Jovanovich.

Sereno, Joan A. and Allard Jongman (1997). Processing of English inflectional morphology. *Memory and Cognition* 25(4): 425–437.

Stampe, David (1969). The acquisition of phonetic representation. In Robert I. Binnick, Alice Davidson, Georgia M. Green and Jerry L. Morgan (eds.), *Papers from CLS 5*, Chicago: Chicago Linguistic Society, pp. 433–444.

Stampe, David (1979). *A dissertation on Natural Phonology*. New York: Garland Pub.

Stampe, David (1987). On phonological representations. In Wolfgang U. Dressler, Hans C. Luschützky, Oskar E. Pfeiffer and John P. Rennison (eds.), *Phonologica 1984: Proceedings of the Fifth International Phonology Meeting, Eisenstadt, 1984*, Cambridge: Cambridge University Press, pp. 287–299.

Taylor, John R. (2003). *Cognitive Grammar*. Oxford: Oxford University Press.

Taylor, John R. (2006). Where do phonemes come from? a view from the bottom. *International Journal of English Studies* 6(2): 19–54.

Tomasello, Michael (2003). Introduction: Some surprises for psychologists. In Michael Tomasello (ed.), *The new psychology of language: Cognitive and functional approaches to language structure*, Mahwah, New Jersey and London: Lawrence Erlbaum Associates, vol. 2, pp. 1–14.

Välimaa-Blum, Riitta (1998). What is deleted in spontaneous Finnish: Segmental interaction with word stress, vowel harmony and moras? In *Proceedings of ESCA workshop Sound Patterns of Spontaneous Speech—Production and Perception (SpoSS), CNRS/Laboratoire Parole et Langage*, Aix en Provence, pp. 47–50.

Välimaa-Blum, Riitta (2005). *Cognitive Phonology in Construction Grammar: Analytic tools for students of English.* Berlin: Mouton de Gruyter.

Välimaa-Blum, Riitta (2008). Do speakers know negative phonototactic constraints in usage-based grammars? Poster at the conference on Language, Communication and Cognition, University of Brighton, August 4–7, 2008.

Välimaa-Blum, Riitta (in press). The phoneme in Cognitive Phonology: Episodic memories of both meaningful and meaningless units? *CogniTextes* 1(2).