To appear in: I. Lasser (ed.) *The Process of Language Acquisition*. Berlin: Peter Lang Verlag.

Coping with phonological variation in early lexical acquisition[*]

Sharon Peperkamp[†‡] and Emmanuel Dupoux[†]
[†]Laboratoire de Sciences Cognitives et Psycholinguistique, EHESS-CNRS, Paris
[‡]Université de Paris 8

1. Introduction

Mapping word forms onto their corresponding meanings is one of the most complex tasks that young infants acquiring their native language have to perform. This is due to the fact that an utterance can refer to many different aspects of a scene, a problem known as referential ambiguity (Quine 1960). An even more basic problem, though, is that it is not easy to find word forms to start with. In fact, the speech waveform is continuous, and word boundaries are not readily available. Moreover, words often surface with different phonetic forms due to the application of postlexical phonological processes; that is, surface word forms exhibit what we call *phonological variation*.

Most models of lexical acquisition assume that infants can somehow extract unique word forms out of the speech stream before they acquire the meaning of words (e.g. Siskind 1996). Hence, they propose a solution to the problem of referential ambiguity, thereby assuming that the problems of finding word boundaries and undoing phonological variation have already been solved. There is evidence that infants can indeed find word boundaries before they have a lexicon (Jusczyk & Aslin 1995). By contrast, virtually nothing is known concerning the question of how prelexical infants deal with phonological variation. In this paper, we will examine if and how infants that do not have a lexicon might undo phonological variation, i.e. deduce which phonological processes apply and infer unique underlying word forms that will constitute lexical entries.

The various intricacies of phonological variation for lexical acquisition can be illustrated within a single language, i.e. Korean. First, consider the allophonic rule of obstruent voicing. This rule voices plain obstruents that occur between two voiced segments, as illustrated in (1); voiced obstruents do not otherwise occur in Korean.

(1)	a.	[kʰɯn**b**aŋ]	‘a big room’
	b.	[jʌp**p**aŋ]	‘a side room’

Obstruent voicing applies in (1a) but fails to apply in (1b), the phonological context not being met in the latter case. The rule thus introduces variation of surface word forms. An infant that segments [baŋ] out of [kʰɯnbaŋ] and [paŋ] out of [jʌppaŋ] might incorrectly entertain the hypothesis that [baŋ] and [paŋ] are two different lexical items. Indeed, quite often, words that occur in the same syntactic position differ only minimally. For instance, Korean has a phonemic aspiration contrast in stop consonants, as shown in (2).[1]

---

[*] Research for this paper was supported by a Fyssen grant to the first author. We would like to thank Hyouk-Keun Kim for help with the Korean examples.
Address for correspondance: Laboratoire de Sciences Cognitives et Psycholinguistique, EHESS-CNRS; 54 Bd. Raspail; 75270 Paris Cedex 6; France
E-mail: {sharon, dupoux}@lscp.ehess.fr.; Web: http://www.ehess.fr/centres/lscp/persons/dupoux
[1] We use a phonetic rather than a phonological representation, since we represent the linguistic input that infants deal with.

(2) a. [**p**i] 'rain'
b. [**pʰ**i] 'blood'

Thus, there is no *a priori* reason why (1) should be a case of allophonic variation and (2) a case of phonemic contrast.[2]

Second, Korean has a rule that assimilates stops to an immediately following nasal consonant, as illustrated in (3); the rule applies in (3a) but not in (3b).

(3) a. [pa**m**noanni] 'did you put the rice?'
b. [pa**p**hænni] 'did you do (i.e. cook) the rice?'

This rule similarly introduces surface variation of word forms. However, the situation is even more complex than in (1), since the resulting assimilated segments are themselves phonemic. Again, an infant that segments [pam] out of [pamnoanni] and [pap] out of [paphænni] might incorrectly entertain the hypothesis that [pam] and [pap] are two different lexical items.

The question we would like to address, then, is how infants might find out that (1) and (3) present instances of surface variation due to the application of a phonological rule, whereas (2) presents a genuine phonemic contrast. In the former cases, the infant should construct a single lexical item, for /paŋ/ in (1) and for /pap/ in (3), while in the latter case, the infant should construct two separate lexical items, for /pi/ and /pʰi/, respectively.

Not all postlexical phonology introduces variation of word forms. Consider, for instance, Korean coda neutralization, illustrated in (4).

(4) a. /tʃipʰ/ [tʃi**p**] ~ [tʃi**pʰ**i] 'straw - straw$_{\mathrm{NOM}}$'
b. /tʃip/ [tʃi**p**] ~ [tʃi**b**i] 'house - house$_{\mathrm{NOM}}$'

The contrast between aspirated and plain stops is neutralized in coda position. It is only in the suffixed forms, in which the stop consonants are in onset position, that this underlying contrast is manifest. In this case, there is no surface variation of word forms; the non-suffixed words always surface as [tʃip]. Rather, variation is present in the phonological realization of roots. Under the assumption that infants initially build lexical entries for word forms only, this type of rule does not pose a problem for early lexical acquisition.

In this paper, we will argue that before learning the meaning of words, infants have sufficient information to undo phonological variation of word forms as illustrated in (1) and (3), that is, deduce the rules and find the underlying forms. By contrast, in the majority of cases, they cannot find the underlying forms of words that present phonological variation at the level of morphemes as illustrated in (4), although they can deduce which rule applies.

As to phonological variation of word forms, we will propose a learning mechanism that deduces which rule applies and infers underlying phonemes and word forms. This mechanism is based on an examination of the distribution of either surface *segments* or surface *word forms*. The distribution of segments will be shown to provide sufficient information in the case of allophonic rules, i.e. rules that involves segments that do not otherwise occur in the language; the distribution of segments that are introduced by this type of rule is complementary to that of segments that are the direct phonetic realization of certain phonemes. For instance, in the example of Korean obstruent voicing in (1), voiced and unvoiced obstruents have complementary distributions. This allows infants to deduce which rule applies and to infer the underlying phonemes. The distribution of word forms will be shown to be necessary in cases in which all surface segments have a phonemic status in the language. In particular, infants can make use of the fact that certain word forms - i.e. the ones

[2] In fact, the opposite pattern is found in English; that is, English has a phonemic voicing contrast in stop consonants, whereas aspiration is allophonic.

that have undergone the rule - fail to occur at the left or right edge of certain phrasal constituents, where the context for application of the rule is never met. For instance, the Korean surface word form [pam], derived from /pap/ by nasal assimilation (see (3)), fails to occur at the end of utterances. Infants can make use of this type of distributional facts to deduce which rule applies and to infer the underlying word forms.

As to phonological variation of morphemes, we will show that although the rules that introduce this type of variation can be deduced in the absence of semantic knowledge, no underlying forms can be inferred. These forms, in fact, can be found only if morphological alternations are taken into account. For instance, whether the Korean surface word form [tʃip] in (4) corresponds to underlying /tʃip^h/ or /tʃip/ can be decided only if it can be compared to a morphologically related form. Presumably, infants that do not know the meaning of words do not have access to such a strategy; consequently, they will initially build lexical entries for surface rather than underlying word forms.

The outline of this paper is as follows. In section 2, we give an overview of experimental data concerning the early stages of phonological and lexical acquisition. We delimit the scope of our proposal in section 3. We then turn to phonological variation of word forms in section 4, and lay out our proposal concerning the way infants might undo this type of variation in the absence of word meanings. In section 5, we consider phonological variation of morphemes, and show that in the majority of cases, it cannot be completely undone without semantic knowledge. Finally, we present our conclusions in section 6.

2. Early phonological and lexical acquisition

Much experimental evidence has been gathered concerning infants' acquisition of phonological properties of their native language during the first year of life. Likewise, the onset of lexical acquisition has been the topic of experimental investigation. In this section, we mention some results that are of direct importance to the issues we address in this paper. For an overview of experimental techniques used in infant speech perception research, assessing infants' discrimination capacities as well as their listening preferences, we refer to Polka, Jusczyk & Rvachew (1995).

First of all, infants are born with the capacity to discriminate all possible vowel and consonant contrasts, regardless of whether they occur in their native language (for an overview of research, see Werker 1991). It has been shown by Kuhl *et al*. (1992) that at 6 months of age, infants react specifically to vowels that are prototypic in their maternal language. Polka & Werker (1994) found that at the same age, infants lose the ability to discriminate non-native vowel contrasts. Together, these results show that 6-month-old infants already know the vowels that occur in their language. The consonantal inventory is acquired somewhat later. That is, it has been shown that between 10 and 12 months, infants lose the ability to discriminate non-native consonantal contrasts, suggesting that they know the consonants that occur in their language (Werker & Tees 1984a).

As to infants' sensitivity to phonotactic properties, there is evidence that it equally arises during the first year of life. Several studies report on infants' listening preferences at 9 months. First of all, at this age, infants prefer to listen to phonotactically legal words rather than to illegal ones (Friederici & Wessels 1993); second, they prefer to listen to non-words with a high-probability phonotactic pattern rather than those with a low-probability phonotactic pattern (Jusczyk, Luce & Charles-Luce 1994); finally, they prefer to listen to unfamiliar words of their native language than to words of a foreign language if the latter violate the phonotactics of their native language (Jusczyk *et al*. 1993).

During the first year of life, infants develop a sensitivity to increasingly smaller phrasal units. First, Hirsh-Pasek *et al*. (1987) showed that at 4½ months, infants listen longer to passages in which pauses were inserted at clause boundaries than at passages in which pauses were inserted within clauses. Second, using the same pause insertion technique, Jusczyk *et al*. (1992) showed that at 9 months, infants are sensitive to phrase boundaries. Finally, Myers *et al*. (1996) found that 11-month-old infants are sensitive to word boundaries. They showed that infants listen longer to passages containing pauses at word boundaries than to passages containing pauses within words. These results hold when the crucial words are unfamiliar, infrequent and not repeated within the passage.

The onset of word segmentation has been reported to lie at 7½ months. Jusczyk & Aslin (1995) found that at this age, infants listen longer to passages containing a word to which they are habituated than to passages that do not contain such a word. The same results are obtained if infants are habituated to passages containing several instances of certain words and tested on these words in isolation. Thus, infants listen longer to words that are contained in the passages they heard previously than to words that are not contained in the passages. Moreover, words appear to be stored in a detailed phonetic representation. For instance, when trained on *cup*, infants show no recognition of *tup*, which differs only as far as the place of articulation of the first segment is concerned.

There is recent experimental evidence that at 10½ months of age, infants recognize the function words of their language. In particular, Shady, Jusczyk & Gerken (1988) reported that at this age, infants listen shorter to passages in which function words are replaced by non-words having the same phonological properties. By contrast, they do not listen shorter if content words are replaced by non-words having the same phonological properties. This suggests that infants not only make a distinction between function words and content words, but also recognize the actual function words of their language.

Various studies have addressed the question as to when the compilation of a lexicon of content words begins. On the basis of comprehension tests as well as observational data in mothers' diary notes, Benedict (1979) reported that 10-month-olds comprehend around 10 words; this figure grows to around 40 at 12 month, and to 100 or more at 16 months. Of course, mothers' diaries may not reflect very accurately the infant's comprehension lexicon. Recent experimental work, however, is consistent with this report. For instance, Hallé & Boysson-Bardies (1994) found that at 10 months of age, infants prefer to listen to a list of 12 familiar rather than to a list of 12 unfamiliar words.

To sum up, much experimental evidence has been gathered regarding early phonological and lexical acquisition. What is actually lacking, though, is an investigation of the acquisition of phonological alternations. In the remaining part of this paper, we will be concerned with the question as to which alternations might be learned without any semantic knowledge. Experimental studies will, of course, be necessary to put our proposals to test.

3. Scope of the proposal

Phonological systems are notoriously complex, in that they combine various rules that interact with one another. We do not propose to solve the entire problem of undoing phonological variation in early language acquisition. Rather, we make several simplifying assumptions in order to evaluate the learnability of fragments of the system of phonological variation and make predictions that can ultimately be tested in infants. Specifically, we consider an idealization of human language that has the following properties.

(5) *Assumptions on linguistic input*:

a. There is no rule interaction; in particular, each segment can be involved in at most one phonological alternation.
b. There are no rules that delete or insert segments.
c. Assimilation rules spread onto an immediately adjacent segment; in other words, there are no long-distance assimilations.

We will implement these assumptions by considering isolated rules taken from real languages.

Concerning infants' innate linguistic knowledge, we make the following assumptions.

(6) *Assumptions on the initial state*:

a. Infants have innate knowledge of two levels of representation in phonology, i.e. an abstract phonological representation and a concrete phonetic representation, with the latter being derived from the former by the smallest possible system of maximally simple rules.
b. Infants have innate knowledge of distinctive features as the primitives of phonological representations, as well as of the feature structure of segments[3] and the types of rules (i.e. delinking and spreading) that may act upon them.
c. Infants have innate knowledge of the hierarchical organization of phonological constituents above the segmental level.[4]

Furthermore, we will exploit the following linguistic abilities that infants acquire during the first year of life, as evidenced by the experimental data discussed in section 2.

(7) *Critical linguistic abilities of infants*:

a. Infants are sensitive to the boundaries of phonological phrases, intonational phrases, and utterances.
b. Infants have a segmental inventory.
c. Infants are sensitive to phonotactics.[5]
d. Infants segment speech into separate surface word forms.

We assume the segmental inventory in (7b) to be phonetic; that is, infants initially make no distinction between segments that occur as allophones only and those that are phonemic.

4. Surface variation of word forms

In this section, we will consider rules that introduce surface variation of word forms, i.e. rules whose structural description makes reference to properties of an adjacent word or to a prosodic boundary. Examples of the former case are assimilation rules that apply across words; the context for assimilation not being always met, words that contain a target for assimilation surface in both an assimilated and an unassimilated form. In the latter case, the rule applies at an edge of a phrasal phonological domain, or, conversely, is blocked at the edge of such a domain; again, the context for the rule will not always be met, hence words containing a target can surface in both an altered and an unaltered form.

We restrict our examination of rules that introduce surface variation of word forms to those that apply at word edges.[6] Our aim is to show that phonological variation at word edges

---

[3] For the purposes of this paper, we do not need to assume a particular version of feature geometry.
[4] We assume the version of the prosodic hierarchy proposed by Nespor & Vogel (1986). The phrasal constituents that are relevant to our research are the phonological phrase, the intonational phrase, and the utterance.
[5] See Hayes (to appear) for an optimality-theoretic approach to infants' acquisition of phonotactics.

can always be undone without any semantic knowledge. Crucially, in all cases, one of the surface word forms corresponds to the underlying form. The infants task is thus, firstly, to find out which rule applies, and, secondly, to determine which of the surface forms corresponds to the underlying form. It should be noted that the underlying form of a word is not necessarily identical to its citation form. For instance, Turkish liquid devoicing, to be discussed in section 4.1.2, applies word-finally at the end of intonational phrases; words pronounced in isolation form an intonational phrase and thus undergo the rule. Therefore, a strategy by which infants attend to one-word utterances, taking these as underlying forms, would be erroneous.[7]

We divide the rules that introduce surface variation of word forms into two classes, i.e. allophonic and non-allophonic ones, and deal with these types in section 4.1. and 4.2, respectively. We will show that undoing allophonic variation is relatively straightforward compared to undoing non-allophonic variation.

4.1 Allophonic variation

Allophonic variation results from rules that introduce segments that do not otherwise occur in the language. The distribution of these segments is thus complementary to that of segments that are phonemic. We will show that the presence of complementary distributions suffices to undo allophonic variation without semantic knowledge. We will examine three cases, Greek /s/-voicing (4.1.1), Turkish liquid devoicing (4.1.2), and Sanskrit *visarga* (4.1.3).

4.1.1 Greek /s/-voicing

Consider the following Greek data. In Greek, [z] is not phonemic, but rather occurs as an allophone of /s/ before voiced consonants. The domain of this rule of /s/-voicing is the intonational phrase (Nespor & Vogel 1986). The rule is illustrated in (8).

(8)	a.	o patera[s]	'the father'
	b.	o patera[z] mas	'our father'
	c.	o patera[s], mu fenete, ine eksipnos	'the father, it seems to me, is bright'

Hence, /s/-voicing applies between words if and only if they are in the same intonational phrase (8b *vs*. 8c). Within intonational phrases, then, the following complementary distributions can be observed:

(9) *Complementary distributions within intonational phrases*:
[z] occurs before voiced consonants only.
[s] occurs everywhere except before voiced consonants (i.e. before vowels and unvoiced consonants, and at the end).

Given these complementary distributions, one of the segments must be present underlyingly and the other one must be an allophone derived by a phonological rule. Under the assumption

[6] Surface variation of word forms can indeed also be found word-internally. For instance, Kimatuumbi has a rule that shortens long vowels in certain words that are non-final within the phonological phrase; this rule applies regardless of the position of the long vowel in the word (Odden 1987). Similarly, tone sandhi rules operate on a vocalic tier, and may thus affect word-internal vowels. We are hence not concerned with this type of rules.

[7] This strategy has several other flaws. For instance, infant directed speech does not necessarily contain many one-word utterances (Aslin *et al*. 1996; Van de Weijer 1999), and it is unclear how infants could distinguish between one-word and multi-word utterances (Christophe *et al*. 1994).

that phonological systems are maximally simple, [z] must be the derived segment, since a rule changing /s/ to [z] is more straightforward than one changing /z/ to [s]. This is due to the fact that the environment in which [z] occurs is more homogeneous than the one in which [s] occurs. Hence, it can be concluded that there is a rule that changes /s/ into [z] before a voiced consonant that is in the same intonational phrase.

The conclusion concerning the allophonic variation can be drawn without word segmentation being available; all that is needed are segmental prototypes and sensitivity to intonational phrases. We therefore predict that once infants begin to segment words and build a recognition lexicon, they will store *pateras* only once. That is, the surface variation between *patera[s]* and *patera[z]* that results from the word-final application of /s/-voicing in, for instance, (8b) does not lead infants to build two separate entries. By the time they begin semantic acquisition, i.e. map word forms onto their corresponding meaning, the phonological variation has already been undone.

4.1.2 Turkish liquid devoicing

A second example comes from Turkish. This language has a rule of liquid devoicing that applies at the end of intonational phrases (Kaisse 1986). Turkish has two liquid consonants, [r] and [l]; an example with the former is shown in (10).

(10)	a.	[ahɨ**r** aldɨ]	'he bought a stable'
	b.	[ačilinja ahɨ**r̥**, ičeri girebiliriz]	'when the stable opens, we can go in'

Within intonational phrases, voiced and unvoiced liquids have complementary distributions:

(11) *Complementary distributions within intonational phrases*:
[r̥] and [l̥] occur only at the end.
[r] and [l] occur everywhere except at the end.

These complementary distributions are in some sense the opposite of what is found with Greek /s/-voicing, since the rule applies at the edge of a domain rather than in the middle. The logic from the acquisitional point of view, however, is the same. That is, given these complementary distributions, it can be inferred that devoiced liquids are allophones of underlying voiced liquids, and that there is a rule that devoices /l/ and /r/ at the end of intonational phrases. As to the mapping of the allophones onto their corresponding phonemes, liquid devoicing consists of delinking of the feature voice; given a universal feature geometry, [r̥] and [l̥] must therefore be allophones of /r/ and /l/, respectively, rather than *vice versa*.

To conclude, we predict that infants deduce the rule of liquid devoicing and infer the phonemes corresponding to the two derived allophones without any semantic knowledge. Once they start to segment speech and build a lexicon, they will only construct lexical items for word forms ending in a voiced liquid.

4.1.3 Sanskrit *visarga*

The final example of an allophonic rule we would like to discuss is Sanskrit *visarga*, which turns utterance-final /r/ and /s/ into a voiceless breathing, called *visarga* (Selkirk 1978). Examples are given in (12).

(12)			*utterance-medial*	*utterance-final*	
	a.	deva/s/	deva[s]	deva[h̥]	'god'
	b.	puna/r/	puna[r]	puna[h̥]	'again'

This rule differs from the Greek and the Turkish ones discussed above in that the allophony is many-to-one rather than one-to-one; hence, it neutralizes the contrast between [s] and [r] at utterance endings. The complementary distributions that can be observed at the surface are stated in (13).[8]

(13) *Complementary distributions within utterances*:
[h̥] occurs only at the end.
[r] and [s] occur everywhere except at the end.

On the basis of these distributions it can be concluded that [h̥] is an allophone of both /r/ and /s/ and that there is a rule that changes /r/ and /s/ into [h̥] at the end of utterances. By contrast, it cannot be decided for a given occurrence of [h̥] whether it derives from an underlying /r/ or /s/. This is not a problem for lexical acquisition, though. In fact, once speech is segmented into separate word forms, words ending in [h̥] can simply be deleted from the list of words that need to be mapped onto a meaning, since the corresponding underlying word forms will necessarily be segmented as well. The latter forms, then, will be the ones that should be mapped onto a meaning.

We conclude that infants coping with this type of allophonic variation can deduce the rule without any semantic knowledge. As to the replacement of a given instance of an allophone by its corresponding phoneme, word segmentation is required. Given that word segmentation necessarily precedes lexical acquisition, infants can from the start correctly build lexical entries for underlying word forms only.

4.2 Non-allophonic variation

We now turn to variation of surface word forms introduced by rules that yield segments with a phonemic status in the language. These rule are not involved in a pattern of complementary distributions. We will propose a learning mechanism based on the comparison of complete word forms that enables to deduce the rule and infer the underlying word forms without semantic knowledge. In addition to those mentioned in (5) in section 3, we will make use of the following assumption on the linguistic input.

(14) *Additional assumption on the linguistic input*:
Lexical words have a non-zero probability of occurrence in both medial and final position within phonological phrases, intonational phrases, and utterances.[9]

We also introduce an additional linguistic ability of infants, acquired during the first year of life.

(15) *Additional critical linguistic ability of infants*:
Infants make a distinction between function words and lexical words.

We will examine two cases, Dutch nasal assimilation (4.2.1) and French voice assimilation (4.2.2).

---

[8] The situation we depict is oversimplified, in that we abstract away from at least two additional facts concerning /s/. First, word-final /s/ assimilates in place of articulation to a following coronal stop; second, word-final /s/ optionally assimilates to any other following segment, and if assimilation does not apply, /s/ turns into [h̥] (cf. Kiparsky 1973).

[9] For the languages considered in this section, i.e. Dutch and French, this is likely to be true.

4.2.1 Dutch nasal assimilation

Dutch has three nasal consonants, /n/, /m/ and /ŋ/. Word-finally, the coronal nasal /n/ assimilates in place of articulation to a following non-coronal consonant (16a); by contrast, /m/ and /ŋ/ never assimilate to a following consonant (16bc).[10]

(16)	a.	tien	[ti**n**]	'ten'
		tien bomen	[ti**m**bomə]	'ten trees'
		tien kamers	[ti**ŋ**kamərs]	'ten rooms'
	b.	stom	[stɔ**m**]	'stupid'
		stom dier	[stɔ**m**dir] / *[stɔ**n**dir]	'stupid animal'
		stom kind	[stɔ**m**kɪnt] / *[stɔ**ŋ**kɪnt]	'stupid child'
	c.	jong	[jɔ**ŋ**]	'young'
		jong dier	[jɔ**ŋ**dir] / *[jɔ**n**dir]	'young animal'
		jong paard	[jɔ**ŋ**part] / *[jɔ**m**part]	'young horse'

Crucially, [n] on the one hand and [m] and [ŋ] on the other hand do not have complementary distributions, since all three nasal consonants are phonemic. Therefore, the inference that in (16a), [m] and [ŋ] are surface manifestations of /n/ cannot be made.

According to Booij (1995), nasal assimilation applies within the domain of the intonational phrase. The following surface regularity, then, appears to hold:

(17) *Surface regularity within intonational phrases*:
There are no clusters consisting of [n] followed by a non-coronal consonant.

This regularity is compatible with two hypotheses. The first one states that certain occurrences of [m] and [ŋ] before labial and velar consonants, respectively, are surface manifestations of /n/, due to regressive nasal assimilation. According to the second hypothesis, certain non-nasal coronal consonants following [n] are surface manifestations of non-coronal consonants, due to progressive assimilation.

We propose that the correct, regressive, assimilation rule can be found without semantic knowledge on the basis of a comparison of complete word forms that occur in medial and final position, respectively, within the intonational phrase. For the sample data in (16a), two lists including the surface word forms given in (18) should be compiled.[11]

(18) *Distribution of surface word forms*:

segmented intonational phrase medially:	[tin] - [tim] - [tiŋ]	- [stɔm] - [jɔŋ]
segmented intonational phrase finally:	[tin] -	- [stɔm] - [jɔŋ]

Note that certain forms, i.e. [tim] and [tiŋ], occur only in medial position. Given the assumption that all lexical items have a non-zero probability to occur at the end of intonational phrases, these forms must be phonological variants of some forms that occur in final position. Moreover, it can be observed that these variants have something in common, i.e. they end in a non-coronal nasal consonant. Coupled with the observation in (17) that there are no clusters of coronal [n] plus a non-coronal consonant, this licenses the hypothesis that [tim] and [tiŋ] are both assimilated forms of underlying /tin/. For [stɔm] and [jɔŋ], such an inference is not drawn, since they occur in both medial and final position. Finally, it can be

[10] We abstract away from two additional facts. First, /n/ equally assimilates to a following labiodental or palatal consonant, yielding the allophones [ɱ] and [ɲ], respectively, that do not occur as phonemes in Dutch; second, within words, all nasal consonants assimilate in place of articulation to a following obstruent (cf. Booij 1995).

[11] That infants can perform this task follows from the joint presence of the linguistic abilities mentioned in (7a) and (7d) in section 3..

checked that, in accordance with this hypothesis, [tim] and [tiŋ] only occur before words beginning with a labial or a velar consonant, respectively, whereas [stɔm] and [jɔŋ] occur everywhere.

Hence, the rule of nasal assimilation can be deduced and the underlying word forms can be inferred without semantic knowledge on the basis of a comparison of surface word forms that occur in medial and final position, respectively, within the intonational phrase. Three caveats, however, are still in order.

First, the data set with which we illustrated the mechanism is extremely limited; a real-life sample would of course be much bigger. Specifically, it would include function words, some of which - e.g. articles - never occur at the end of intonational phrases. One might thus raise the question how to prevent the incorrect inference that function words that occur in medial position only are phonological variants of some word forms that occur in final position. We suggest that function words are processed as a separate class, which is compatible with the fact that infants show an early ability to distinguish between function words and content words.

Second, a few words should be said about homophony. Given that Dutch nasal assimilation is non-allophonic, an assimilated form can be homophonous to another word in the language. An example is given in (19).

(19)	a.	Han	[hɑ**n**]	'id. (proper name)'
		ham	[hɑ**m**]	'id.'
		hang	[hɑ**ŋ**]	'bent'
	b.	Han bidt	[hɑ**m**bɪt]	'Han prays'
		Han koopt	[hɑ**ŋ**kopt]	'Han buys'

The three lexical items in (19a) differ only as far as the place feature of their final nasal consonant is concerned. In (19b), it is shown that due to assimilation, the first lexical item can have the same surface form as the two other ones. Given that the latter two can occur in final position within the intonational phrase, the distributions of word forms in medial and final position within the intonational phrase are completely identical. Therefore, the inference that [hɑm] and [hɑŋ] might be assimilated forms of underlying /hɑn/ cannot be made. For the case at hand, this is not a problem, since separate lexical entries should indeed be built not only for [hɑn] but also for [hɑm] and [hɑŋ]. However, if there is much surface homophony, there will be few word forms ending in [m] or [ŋ] that are found in medial position only. Their absence in final position, then, might be interpreted as an accidental gap rather than as a systematic fact. Consequently, the general hypothesis according to which such forms have undergone nasal assimilation might not be made. At present, we do not know the percentage of homophony of surface word forms in Dutch. An actual learning algorithm, however, should be capable of dealing with this type of noise in the signal.

The third caveat concerns the fact that we have tacitly assumed that nasal assimilation is obligatory. Booij (1995), however, describes this rule as being optional. More realistic data, then, would look like the ones in (20).

(20) *Optional assimilation*

a.	tien	[ti**n**]			'ten'
b.	tien dieren	[ti**n**dirə]			'ten animals'
c.	tien bomen	[ti**m**bomə]	-	[ti**n**bomə]	'ten trees'
d.	tien kamers	[ti**ŋ**kamərs]	-	[ti**n**kamərs]	'ten rooms'

Two observations regarding these data should be made. On the one hand, contrary to what is found with obligatory assimilation, the coronal nasal [n] is not banned before non-coronal consonants (20cd). On the other hand, on a par with obligatory assimilation, the surface forms

[tim] and [tiŋ] fail to occur in final position within the intonational phrase. In other words, the surface regularity stated in (17) does not hold, but the distributions of word forms in medial and final position in the intonational phrase are identical to those in (18). Since our proposal for a learning mechanism relies on both a surface regularity and a differential distribution, the case of optional rules seems problematic. We propose two possibilities to solve the problem.

The first solution is that it might be possible to devise an alternative learning mechanism that does not rely on the presence of a surface regularity. In fact, in our proposal this regularity is mainly used to restrict the set of rules that may be involved and, hence, to accelerate the acquisition process. Given that the number of possible rules is finite, this restriction might not be essential. The second solution is that a surface regularity might be found if one or several of the following factors are taken into account. First, assimilation is more likely to apply in casual speech than in careful speech. Second, assimilation is more likely to apply in fast speech than in speech at a normal or a slow rate. Third, assimilation might be optional in the intonational phrase but obligatory within the phonological phrase. Hence, if speech style and rate as well as the size of the prosodic domain are taken into account, restrictions on the occurrence of [n] might be noted, and the hypothesis that there is a nasal assimilation rule might be raised.

To sum up, variation of surface word forms due to the rule of nasal assimilation can be undone without semantic knowledge. We predict that infants can deduce this rule and infer the underlying word forms once they can segment continuous speech into separate word forms and store these forms into a recognition lexicon. As to the optionality of the rule, it would be interesting to see whether in child-directed speech, assimilation is more frequent than in adult-directed speech.

The case of Dutch nasal assimilation is relatively straightforward for two reasons. First, the triggers (i.e. non-coronal consonants) and the target (i.e. the coronal nasal /n/) form completely distinct sets. Second, the resulting segments (i.e. [m] and [ŋ]) are not targets for assimilation themselves. In the next section, we will examine an assimilation rule in which these conditions do not hold.

4.2.2 French voice assimilation

Consider the data in (21), illustrating the rule of French voice assimilation.

(21)	a.	robe	[ʀɔ**b**]	‘dress’
		robe jaune	[ʀɔ**b**ʒo:n]	‘yellow dress’
		robe sale	[ʀɔ**p**sal]	‘dirty dress’
	b.	patte	[pa**t**]	‘paw’
		patte sale	[pa**t**sal]	‘dirty paw’
		patte jaune	[pa**d**ʒo:n]	‘yellow paw’

Both triggers and targets are obstruents. As can be seen from these data, voice assimilation is regressive and spreads both [+voice] and [-voice]. In the absence of the word meanings, however, only the following surface regularity will be observable:[12]

(22) *Surface regularity within utterances*:
Obstruent clusters agree in voicing.

---

[12] We do not know what the domain of assimilation is. For the sake of simplicity, we assume that assimilation applies throughout the utterance. Given that infants are sensitive to utterances, intonational phrases and phonological phrases alike (see section 2), the domain of the rule is irrelevant to our argument. That is, if the domain turns out to be, for instance, the intonational phrase, one should read ‘intonational phrase’ instead of ‘utterance’ in what follows.

This observation leads to the hypothesis that there is a voice assimilation rule. Assimilation, though, can be of different types. In fact, assimilation rules universally involve two parameters, stated in (23).

(23) *Universal parameters in assimilation*
a. direction of assimilation: regressive, progressive, or bi-directional
b. spreading value: either marked or both marked and unmarked

We argue that the values of these parameters can be set in the absence of semantics, as follows.

First of all, by crossing the two parameters in (23), we obtain six types of assimilation, one of which, however, is logically impossible. That is, bi-directional assimilation of both the marked and the unmarked value is ruled out, since spreading could be either to the left or to the right, yielding different results.[13] As a first step towards setting the values of the two assimilation parameters, it can be deduced which types of assimilation are in accordance with the surface regularity that obstruent clusters agree in voicing. That is, assimilation types that yield disagreeing clusters on the surface are immediately ruled out. Table 1 shows for each assimilation type whether the two disagreeing clusters, abbreviated as [+ -] and [- +], can surface or not. In the latter case, they are preceded by '*'.

spreading value(s)	direction of assimilation		
	regressive	progressive	bi-directional
[+voice]	[+ -] *[- +]	*[+ -] [- +]	*[+ -] *[- +]
[+voice],[-voice]	*[+ -] *[- +]	*[+ -] *[- +]	

Table 1: Disagreeing clusters in voice assimilation typology

Thus, the following three types of assimilation yield surface patterns in which both disagreeing clusters are ruled out:

(24) a. regressive assimilation; spreading of both [+voice] and [-voice]
b. progressive assimilation; spreading of both [+voice] and [-voice]
c. bi-directional assimilation; spreading of [+voice] only

Next, separate lists can be compiled of surface word forms that occur utterance-medially and those that occur utterance-finally. For the sample data in (21), these lists look like as shown in (25).

(25) *Distribution of surface word forms within utterances*:
occurrence in medial position: [ʀɔb] - [ʀɔp] - [pat] - [pad] - [ʒoːn] - [sal]
occurrence in final position: [ʀɔb] - - [pat] - - [ʒoːn] - [sal]

Certain forms, i.e. [ʀɔp] and [pad], occur in utterance-medial position only. Under the assumption that all lexical items have a non-zero probability to occur utterance-finally, these forms must be phonological variants of some forms that occur in final position. Moreover, it can be observed that these variants have something in common, i.e. they end in an obstruent. This licenses the hypothesis that they are the result of the application of voice assimilation.[14]

[13] For instance, in a cluster consisting of a voiced consonant followed by an unvoiced consonant, leftwards spreading would result in a uniformly unvoiced cluster while rightwards spreading would result in a uniformly voiced cluster.
[14] Of course, in this extremely small data set, the fact that the phonological variants end in an obstruent while no such regularity holds for, say, their initial segments might as well be interpreted as being accidental. A real-life

It is now easy to see that assimilation is regressive. First, given that the phonological variation concerns the final consonant in the word, spreading must be onto this consonant, i.e. from right to left. Hence, progressive assimilation is ruled out. Second, given that the final consonant of the phonological variants can surface as either voiced (as in [pad]) or unvoiced (as in [ʀɔp]), both values of the feature voice must spread. Therefore, bi-directional assimilation, involving spreading of [+voice] only, is ruled out as well. Regressive assimilation, then, with spreading of both feature values, is the only remaining possibility. In other words, [ʀɔp] and [pad] are recognized as phonological variants of /ʀɔb/ and /pat/, respectively.[15]

To sum up, we have shown that phonological variation due to regressive voice assimilation with spreading of both feature values can be undone in the absence of semantics. It appears that variation due to the two other assimilation rules that yield the surface regularity in (22) is harder to undo. Recall that these rules involve progressive spreading of both feature values and bi-directional spreading of [+voice], respectively (see (24)). In order to illustrate how surface variation introduced by these types of assimilation can be undone, we consider two pseudofrench languages that are identical to French except that their voice assimilation rules are progressive with spreading of both values and bi-directional with spreading of [+voice], respectively. With progressive assimilation, spreading is onto word-initial consonants only (26), while with bi-directional assimilation, spreading is onto both word-initial and word-final consonants (27); in the examples, assimilated consonants are in boldface.

(26) *Pseudofrench I: progressive assimilation*

a.	robe sale	[ʀɔb**z**al]	'dirty dress'
	robe jaune	[ʀɔbʒo:n]	'yellow dress'
b.	patte sale	[patsal]	'dirty paw'
	patte jaune	[pat**ʃ**o:n]	'yellow paw'

(27) *Pseudofrench II: bi-directional assimilation*

a.	robe sale	[ʀɔb**z**]ale	'dirty dress'
	robe jaune	[ʀɔbʒo:n]	'yellow dress'
b.	patte sale	[patsal]	'dirty paw'
	patte jaune	[pa**d**ʒo:n]	'yellow paw'

Crucially, if spreading is onto a word-initial consonant, the context for assimilation is never met in utterance-initial position. In other words, the underlying forms surface utterance-initially rather than utterance-finally. Hence, in order to find the differences in the distribution of surface word forms in various positions in the utterance, the focus of attention should be onto utterance beginnings in the case of progressive assimilation and onto both utterance beginnings and utterance endings in the case of bi-directional assimilation. The distributions of surface word forms in initial, medial, and final position in the two pseudo-languages are shown in (28) and (29).

(28) *Pseudofrench I, distribution of surface word forms within utterances*:

occurrence in initial position:	[ʀɔb] - [pat] - [ʒo:n] - - [sal]
occurrence in medial position:	[ʀɔb] - [pat] - [ʒo:n] - [ʃo:n] - [sal] - [zal]
occurrence in final position:	[ʀɔb] - [pat] - [ʒo:n] - [ʃo:n] - [sal] - [zal]

sample, however, would be much bigger; hence the hypothesis that the variation is due to voice assimilation would be raised on a more solid basis.

[15] On a par with the case of Dutch nasal assimilation, three caveats regarding function words, homophony, and optionality, respectively, are in order. See section 4.2.1 for discussion.

(29) *Pseudofrench II, distribution of surface word forms within utterances*:

occurrence in initial position:	[ʀɔb] - [pat] - [pad] - [ʒoːn] - [sal]
occurrence in medial position:	[ʀɔb] - [pat] - [pad] - [ʒoːn] - [sal] - [zal]
occurrence in final position:	[ʀɔb] - [pat] - - [ʒoːn] - [sal] - [zal]

These distributions are misleading, though, in that the great majority of utterances begin with a function word, at least in right-recursive languages such as French. Consequently, a great many lexical words never surface utterance-initially, and the rule cannot be deduced.[16] As a solution to this problem, it might be argued that function words can be stripped off from the beginnings of utterances; the focus of attention, then, is on the first lexical word. However, given that assimilation can apply between a function word and a following lexical word, there would also be assimilated forms in the first lexical position; therefore, there would be no differentiated distributions of word forms found in medial position and those found in the first lexical position.

Alternatively, we propose a way of deducing progressive and bi-directional assimilation that focuses on utterance endings only. It departs from the observation that there are three types of assimilation that are in accordance with the surface regularity, and aims at eliminating two of these. Consider first the distribution of word forms in medial and final position in Pseudofrench I, thus ignoring the initial position altogether. These distributions are completely identical, hence no form can be identified as a phonological variant. Therefore, there is no leftwards spreading, and assimilation can thus be neither regressive nor bi-directional. Assimilation, then, is progressive, and - given the surface regularity - both values spread.

The case of Pseudofrench II is slightly more complicated. If a comparison is made between forms in medial and final position, one form, i.e. [pad], is identified as a phonological variant. In fact, [pad] occurs in medial but not in final position. Concerning this form, it can be observed - given the complementary presence of [pat] - that the phonological variation concerns the final consonant of the word. Hence, the assimilation rule must include leftwards spreading. Moreover, this spreading involves the feature [+voice] only, since there is no phonological variant ending in an unvoiced consonant. Given these two observations, progressive and regressive assimilation with spreading of both feature values are both ruled out. This allows to draw the conclusion that assimilation is bi-directional, with spreading of [+voice] only.

We have shown, then, that the presence of progressive or bi-directional assimilation can be deduced without taking utterance beginnings into account. However, in order to completely undo the phonological variation, underlying word forms should also be inferred. This step appears to be problematic in Pseudofrench I and II. In fact, in Pseudofrench I, all underlying forms surface at utterance beginnings, *modulo* the occurrence of initial function words, while in Pseudofrench II, half of the underlying forms surfaces in this position. Hence, we are back at the problem that the learning mechanism needs information that is obscured by the presence of function words. We would like to suggest a solution that - for reasons of space - we can only sketch here. Essentially, we propose to use a more refined analysis of the contexts in which word forms occur. Indeed, the surface regularity allows to determine the contexts in which assimilation might apply, i.e. preceding or following a word that begins or ends with an obstruent, respectively. Thus, instead of classifying words as occurring in medial or final position within an utterance, one could store whether they are in the context of

[16] One might hope to find a correlation between the direction of syntactic recursion and that of across-word assimilation rules. In particular, if right-recursive languages had only regressive assimilation rules and left-recursive languages had only progressive assimilation rules, the problem of function words obscuring the differential distributions would be reduced to cases of bi-directional assimilation. The correlation, however, does not exist; a counterexample is Dutch, which is syntactically right-recursive but has a progressive voice assimilation rule (Booij 1995).

assimilation or not. That is, two lists of word forms beginning or ending in an obstruent would be compiled; one list would contain all word forms whose obstruent forms a cluster with an initial or final obstruent in an adjacent word form, the other list would contain all remaining word forms. Crucially, the latter list would be the more restricted one, exclusively containing surface word forms that are identical to underlying forms.

To sum up this section, we have shown that surface variation of word forms introduced by three voice assimilation rules that yield a single surface regularity can be undone without semantic knowledge.[17] Analogously to the case of Dutch nasal assimilation in section 4.1, we predict that infants can deduce this type of rules and infer the underlying word forms once they segment the speech stream into separate words and build a recognition lexicon (i.e. store surface word forms).

4.3 Summary

We have argued that surface variation of word forms can be undone without semantic knowledge, regardless of whether the variation is allophonic or non-allophonic. The cases we examined can also be classified according to whether the rules that introduce the variation are neutralizing or non-neutralizing. Notice that non-allophonic rules - i.e. rules are of the type "α becomes β" where β is a phoneme in the language - are necessarily neutralizing; they neutralize the distinction between the phonemes α and β. Allophonic rules, by contrast, can be either neutralizing or non-neutralizing. By crossing the dimensions of allophony and neutralization we thus obtain a three-way classification, shown in Table 2. In this table, we note whether word segmentation is a prerequisite for the deduction of the rule and the inference of underlying forms, respectively. If word segmentation is required, the process is said to be 'difficult'; otherwise, it is said to be 'easy'.

	deduction of the rule	inference of underlying forms
allophonic non-neutralizing		
Greek /s/-voicing;	easy	easy
Turkish liquid devoicing	easy	easy
allophonic neutralizing		
Sanskrit visarga	easy	difficult
non-allophonic neutralizing		
Dutch nasal assimilation;	difficult	difficult
French voice assimilation	difficult	difficult

Table 2: Undoing surface variation of word forms without semantics

Variation due to rules that are allophonic and non-neutralizing, such as Greek /s/-voicing and Turkish liquid devoicing, are the easiest to undo in the absence of semantics; the first property corresponds to exhibiting a pattern of complementary distributions, while the second one assures a one-to-one correspondence between surface forms and underlying forms. Word segmentation, then, does not come into play at all. Variation due to allophonic neutralizing rules, such as Sanskrit *visarga*, is slightly more complex, since the one-to-one correspondence is lacking; word segmentation is therefore required in order to infer the underlying forms. Variation due to non-allophonic (hence, neutralizing) rules, such as Dutch nasal assimilation and French voice assimilation, is the hardest to undo, since there is no pattern of

[17] We do not deal with the two remaining types, involving leftwards and rightwards spreading of the marked feature [+voice]. Note that Dutch nasal assimilation, treated in 4.2.1, is similar to the first one of these; in fact, this rule spreads the marked place features labial and velar to the left.

complementary distributions to begin with. This type of variation, then, can be undone only if word segmentation is available and a recognition lexicon can be constructed.

5. Surface variation of morphemes

As we have seen above, the process of undoing phonological variation consists of two steps. First, the rule that introduces the variation must be deduced. Second, given a surface form, the underlying form must be inferred. In the previous section, we have shown that both tasks can be accomplished without semantics in the case of surface variation of word forms. In this section, we turn to surface variation of morphemes rather than of words, and argue that in the absence of semantics, rules that introduces such variation can always be deduced; inferring underlying forms, by contrast, can be accomplished only if the rule is allophonic and non-neutralizing.

Surface variation of morphemes is introduced by rules that do not make reference to properties of an adjacent word or to a prosodic boundary larger than that of the prosodic word. The variation can appear both word-internally and at a word edge; we will consider examples of both types. We will examine allophonic variation first (5.1), and then turn to non-allophonic variation (5.2).

5.1 Allophonic variation

We will consider three cases of allophonic variation; the first two cases concern non-neutralizing rules, i.e. Japanese affrication (5.1.1) and Italian /s/-voicing (5.1.2), while the third case concerns a neutralizing rule, i.e. Chamorro vowel reduction (5.1.3). We will argue that only allophonic non-neutralizing variation can be completely undone in the absence of semantics.

5.1.1 Japanese affrication

In Japanese, /t/ turns into [t^s] whenever it precedes the vowel [u]; the affricate [t^s] does not otherwise occur in the language[18] (Itô & Mester 1995). This is illustrated in (30).

(30) a.	[ka**t**-anai]	[ka**t^s**-u]	'win$_{NEG}$ - win$_{PRES}$'
b.	[kakikukeko]	[tatʃi**t^s**uteto]	'k-column - t-column'

Given that no word can end in [t], the context for affrication is never met across words. Hence, this allophonic rule introduces surface variation of morphemes rather than of word forms. Moreover, [t] and [t^s] have complementary distributions:

(31) *Complementary distributions*:
[t^s] occurs only before [u].
[t] occurs everywhere except before [u].

Each surface [t^s], then, must correspond to an underlying /t/, and this conclusion can be drawn on the basis of a phonetic representation only, that is without access to word segmentation and semantic knowledge. Therefore, we predict that infants can deduce this rule and infer the

---

[18] We abstract away from the occurrence of [t^s] before other vowels in foreign loans, such as *t^saitogaisuto* 'Zeitgeist'.

underlying phoneme before they begin to segment speech into separate words; they should thus recode each instance of [t^s] as an underlying /t/.

5.1.2 Italian /s/-voicing

As opposed to the case of Japanese affrication, much allophonic variation of morphemes requires word segmentation in order to be undone. In particular, this holds for variation resulting from rules that are blocked across words. Consider, for instance, Italian /s/-voicing. This rule voices /s/ if it precedes a voiced consonant in the same word; the resulting [z] does not otherwise occur in the language (Camilli 1965).[19] This rule thus introduces variation of morphemes, as illustrated for the prefix *dis*- in (32).

(32)	a.	disdire	[di**z**dire]	'to denie'
	b.	disfare	[di**s**fare]	'to undo'

The rule applying within words only, [s] and [z] do not have complementary distributions within a phrasal domain. To see this, consider (33), where the rule applies within a word (33a) but fails to apply in a two-word phrase (33b).

(33)	a.	bislungo	[bi**z**luŋo]	'oblong'
	b.	bis lungo	[bi**s**luŋo]	'long *encore*'

In the absence of word boundaries, only the following surface regularity can be observed.

(34) *Surface regularity*:
[z] occurs before voiced consonants only.

This regularity is compatible with at least two hypotheses: either [z] is an allophone of /s/ that appears before voiced consonants within some domain, or there is a phoneme /z/ that devoices before unvoiced consonants as well as before vowels.

Which of these two hypotheses is correct cannot be decided unless word boundaries can be taken into account. Complementary distributions within words can then be observed:

(35) *Complementary distributions within words*:
[z] occurs before voiced consonants.
[s] occurs everywhere except before voiced consonants.

The possibility that /z/ is a phoneme that devoices before unvoiced consonants as well as before vowels is now ruled out, since it fails to account for the fact that [s] does not occur before voiced consonants.[20] Therefore, [z] must be an allophone of /s/ that appears before voiced consonants only.

To conclude, we predict that infants can undo the surface variation due to /s/-voicing once they have acquired word segmentation.

5.1.3 Chamorro vowel reduction

[19] At least not in central and southern varieties of Italian; in northern varieties, [z] also occurs as an allophone of /s/ between vowels (Nespor & Vogel 1986).
[20] One might argue that this was not a viable option to begin with, given our assumption that each segment is involved in at most one phonological rule. Indeed, both assimilation (before unvoiced consonants) and dissimilation (before vowels) would be involved.

As an example of an allophonic neutralizing rule, consider Chamorro vowel reduction (Topping 1968). In Chamorro, /u/ and /o/ are reduced to [ʊ] in unstressed syllables.[21] This is illustrated in (36).

(36)	a.	[dágʊ]	~	[i dəg**ú**-hù]	'yam – my yam'
	b.	[pécʊ]	~	[i pɪc**ó**-kù]	'chest – my chest'

The allophone [ʊ] does not otherwise occur in the language. Under the assumption that stress is available as a distinctive suprasegmental property of vowels, the following complementary distributions can be observed:

(37) *Complementary distributions*:
[ʊ] occurs without stress only.
[u] and [o] occur under stress only.

Note that - as opposed to all other cases of complementary distributions we have examined so far - these distributions are equally homogeneous. Homogeneity is therefore not available as a cue for deducing which rule applies. However, given that phonological rules are deterministic, [u] and [o] cannot be both derived by rule from a single underlying phoneme /ʊ/. As a consequence, the distributions in (37) are only compatible with the hypothesis that there is a rule of vowel reduction, changing /u/ and /o/ into [ʊ].

We have shown, then, that the rule of vowel reduction can be deduced without word segmentation being available. It is easy to see, though, that even when surface word forms are given, the underlying vowel for a given occurrence of [ʊ] cannot be inferred. In fact, the alternations in (36) introduce variation in the phonological realization of morphemes rather than complete word forms; the words in the first column can be related to those in the second one only on the basis of lexical knowledge, i.e. morphology and semantics. Consequently, we predict that in the phonological representation of words in the mental lexicon, infants will include [ʊ] rather than the underlying full vowel from which it is derived; as morphology becomes available, they can recode the words more abstractly and delete [ʊ] from the phonological representation.

5.2 Non-allophonic variation

We are now left with non-allophonic variation of morphemes, which can only be introduced by neutralizing rules. Using an example from German, we will show that this type of variation cannot be completely undone in the absence of semantics.

5.2.1 German final devoicing

In German, the underlying contrast between voiced and unvoiced obstruents is neutralized in syllable-final position. In that position, only unvoiced obstruents surface. This neutralization is illustrated in (38).

(38)	a.	/bunt/	[bun**t**]	~	[bun.**t**e]	'multi-colored$_{\text{MASC-FEM}}$'
	b.	/bund/	[bun**t**]	~	[bun.**d**e]	'league$_{\text{NOM-DAT}}$'

[21] Chamorro has a six-vowel system. We abstract away from the fact that the remaining four vowels similarly undergo pair-wise reduction to two other allophones.

Unvoiced obstruents being phonemic, their distribution is not complementary to that of voiced obstruents. However, once word segmentation is available, the following surface regularity can be observed:

(39) *Surface regularity within words*:
Voiced obstruents do not occur in final position.[22]

Given that voiced obstruents are found in other positions, this regularity triggers the hypothesis that word-finally, voiced obstruents are devoiced.

The rule can thus be deduced if word segmentation is taken into account. However, given that the rule introduces variation in the phonological realization of morphemes but not in that of complete word forms, the underlying voicing value for a given word-final obstruent cannot be inferred without morphology. In fact, the words in the left-hand column in (38) can be related to those in the right-hand one only on the basis of lexical knowledge, i.e. morphology. We predict, then, that infants will initially store words in their mental lexicon with word-final unvoiced obstruents, regardless of whether these obstruents are underlyingly voiced or voiceless; it is only when morphology and semantics become available that infants can infer the underlying consonants and modify the phonological representation of words accordingly.

5.3 Summary

Table 3 summarizes the results regarding surface variation of morphemes. The rules are classified according to whether they are allophonic and/or neutralizing. In this table, we note whether in the absence of semantics, the rules can be deduced and the underlying forms can be inferred, respectively. As in Table 2, 'easy' stands for a process without word segmentation, while 'difficult' denotes that word segmentation is required.

	deduction of the rule	inference of underlying forms
allophonic non-neutralizing		
Japanese affrication;	easy	easy
Italian /s/-voicing	difficult	difficult
allophonic neutralizing		
Chamorro vowel reduction	easy	impossible
non-allophonic neutralizing		
German final devoicing	difficult	impossible

Table 3: Undoing surface variation of morphemes without semantics

Only surface variation introduced by allophonic non-neutralizing rules, such as Japanese affrication and Italian /s/-voicing, can be completely undone without semantics. The case of Japanese affrication is particularly straightforward, since word segmentation is unnecessary. Variation introduced by allophonic neutralizing or non-allophonic neutralizing rules, such as Chamorro vowel reduction and German final devoicing, respectively, cannot be completely undone without semantics; that is, while the rules can be deduced, the underlying forms cannot be inferred in the absence of semantic knowledge.

[22] We assume that word segmentation precedes syllable segmentation. Hence, the correct generalization that voiced obstruents fail to occur syllable-finally, will be made only at some later stage.

6. Conclusions

In this paper, we have examined surface variation of word forms or morphemes introduced by postlexical phonological rules. We have decomposed the process of undoing phonological variation into two steps, consisting of the deduction of the rule and the inference of underlying phonemes and word forms, respectively. We have argued that a rule can always be deduced in the absence of semantics. As to the inference of underlying phonemes and word forms, we have argued that if the phonological variation concerns word forms it can always be drawn in the absence of semantics; if the variation concerns morphemes only, by contrast, it cannot in general be drawn (the exception being the case of allophonic non-neutralizing rules). In terms of early lexical acquisition, these results boil down to the following generalizations. On the one hand, all phonological variation that complicates the problem of lexical acquisition, i.e. variation of word forms, can be undone by infants before constructing a lexicon. On the other hand, most of the phonological variation that does not interfere with early lexical acquisition, i.e. variation of morphemes, cannot be undone at this stage.[23] Hence, we conclude that surface variation introduced by postlexical phonology does not interfere with early lexical acquisition; that is, infants can find all and only those word forms, i.e. the underlying word forms, that need to be mapped onto a meaning before they actually begin to resolve the form-to-meaning problem.[24]

Of course, we made many simplifying assumptions, and the real task faced by infants is much more complex. For instance, languages generally have more than one postlexical rule, and, more importantly, rules often interact with one another. A consequence of rule interaction is that some rules are opaque, in that there is no surface true generalization pertaining to them. That is, either their output or their context of application is transformed by the application of another rule. Despite the numerous simplifications, our approach allows to make predictions regarding the order in which various types of phonological variation are undone by infants. Specifically, we have shown that for certain types of variation, both the deduction of the rules and the inference of underlying forms are 'easy', in that they can be coped with without word segmentation (see Tables 2 and 3). Hence, these types of variation can be undone by infants within the first year of life. Other types of phonological variation are 'difficult' to undo, in that the inference of underlying forms and often also the deduction of the rule require word segmentation. We expect that infants will only be able to undo these types of variation once they have stored a fair amount of lexical items into a recognition lexicon, that is, presumably only within their second year of life. Finally, some types of phonological variation cannot be completely coped with without semantic knowledge, since the underlying forms cannot be inferred. We predict that infants will not undo these types of variation until they are well advanced into the sound-to-meaning acquisition and begin to decompose words into morphemes.[25] We plan to carry out both experiments with infants and computer simulations to test these predictions.

References

Aslin, R., J. Woodward, N. LaMendola, and T. Bever (1996) Models of word segmentation in fluent maternal speech to infants. In: J. Morgan and K. Demuth (eds.) *Signal to Syntax. Bootstrapping from Speech to Grammar in Early Acquisition*. Mahwah, NJ: LEA, 117-134.

[23] We assume that infants initially build lexical entries for word forms only.

[24] For experimental evidence that lexical items are stored in an underlying representation, at least in adults, see Gaskell & Marslen-Wilson (1996, 1998) and Lahiri & Marslen-Wilson (1991).

[25] In this paper, we have not dealt with lexical phonological rules, but we predict these to be acquired after all postlexical phonology. In fact, lexical rules not only involve alternations at the morphemic level, but they also make reference to specific morphemes and contain lexical exceptions.

Benedict, H. (1979) Early lexical development: comprehension and production. *Journal of Child Language* 6, 183-200.

Booij, G. (1995) *The Phonology of Dutch*. Oxford: Oxford University Press.

Camilli, A. (1965) *Pronunica e grafia dell'italiano*. Firenze: Sansoni.

Christophe, A., E. Dupoux, J. Bertoncini, and J. Mehler (1994) Do infants perceive word boundaries? An empirical study of the bootstrapping of lexical acquisition. *Journal of the Acoustical Society of America* 95, 1570-1580.

Friederici, A. and J. Wessels (1993) Phonotactic knowledge of word boundaries and its use in infant speech perception. *Perception and Psychophysics* 54, 287-295.

Gaskell, M. and W. Marslen-Wilson (1996) Phonological variation and interference in lexical access. *Journal of Experimental Psychology: Human Perception and Performance* 22, 144-158.

Gaskell, M. and W. Marslen-Wilson (1998) Mechanisms of phonological interference in speech perception. *Journal of Experimental Psychology: Human Perception and Performance* 24, 380-396.

Gerken (1996) Phonological and distributional information in syntax acquisition. In: J. Morgan and K. Demuth (eds.) *Signal to Syntax. Bootstrapping from Speech to Grammar in Early Acquisition*. Mahwah, NJ: LEA, 411-425.

Hallé, P. and B. de Boysson-Bardies (1994) Emergence of an early receptive lexicon: infants' recognition of words. *Infant Behavior and Development* 17, 119-129.

Hayes, B. (to appear) Phonological acquisition in optimality theory: the early stages. To appear in: R. Kager, W. Zonneveld, and J. Pater (eds.) *Phonological Acquisition and Typology*.

Hirsh-Pasek, K., D. Kemler-Nelson, P. Jusczyk, K. Wright Cassidy, B. Druss, and L. Kennedy (1987) Clauses are perceptual units for young infants. *Cognition* 26, 269-286.

Itô, J. and A. Mester (1995) Japanese phonology. In: J. Goldsmith (ed.) *The Handbook of Phonological Theory*. Oxford: Blackwell, 817-838.

Jusczyk, P. and R. Aslin (1995) Infants' detection of sound patterns of words in fluent speech. *Cognitive Psychology* 29, 1-23.

Jusczyk, P., A. Friederici, J. Wessels, V. Svenkerud, and A. Jusczyk (1993) Infants' sensitivity to the sound pattern of native language words. *Journal of Memory and Language* 32, 402-420.

Jusczyk, P., K. Hirsh-Pasek, D. Kemler- Nelson, L. Kennedy, A. Woodward, and J. Piwoz (1992) Perception of acoustic correlates of major phrasal units by young infants. *Cognitive Psychology* 24, 252-293.

Jusczyk, P., P. Luce, and J. Charles-Luce. (1994) Infants' sensitivity to phonotactic patterns in the native language. *Journal of Memory and Language* 33, 630-645.

Kaisse, E. (1986) Locating Turkish devoicing. *Proceedings of WCCFL* 5. Stanford Linguistics Association, 119-128.

Kiparsky, P. (1973) "Elsewhere" in phonology. In: S. Anderson and P. Kiparsky (eds.) *A Festschrift for Morris Halle*. New York: Holt, 93-106.

Kuhl, P., K. Williams, F. Lacerda, K. Stevens, and B. Lindblom (1992) Linguistic experience alters phonetic perception in infants by six months of age. *Science* 255, 606-608.

Lahiri, A. and W. Marslen-Wilson (1991) The mental representation of a lexical form: A phonological approach to the recognition lexicon. *Cognition* 38, 245-294.

Myers, J., P. Jusczyk, D. Kemler-Nelson, J. Charles-Luce, A. Woodward, and K. Hirsh-Pasek (1996) Infants' sensitivity to word boundaries in fluent speech. *Journal of Child Language* 23, 1-30.

Nespor, M. and I. Vogel (1986) *Prosodic Phonology*. Dordrecht: Foris.

Odden, D, (1987) Kimatuumbi phrasal phonology. *Phonology* 4, 13-36.

Polka, L., P. Jusczyk, and S. Rvachew (1995) Methods for studying speech perception in infants and children. In: W. Strange (ed.) *Speech Perception and Linguistic Experience: Issues in Cross-language Research*. Baltimore, MD: York Press, 49-89.

Polka, L. and J. Werker (1994) Developmental changes in perception of non-native vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance* 20, 421-435.

Quine, W. (1960) *Word and Object*. Cambridge, MA: MIT Press.

Selkirk, E. (1978) Prosodic domains in phonology: Sanskrit revisited. In: M. Aronoff and M.-L. Kean (eds.) *Juncture*. Saratoga, CA: Anma Libri, 107-129.

Shady, M., P. Jusczyk, and L. Gerken (1988) *Infants' sensitivity to function morphemes*. Paper presented at the 23rd Annual Boston University Conference on Language Development, Boston, MA.

Siskind, J. (1996) A computational study of cross-situational techniques for learning word-to-meaning mappings. *Cognition* 61, 39-91.

Topping, D. (1968) Chamorro vowel harmony. *Oceanic Linguistics* 7, 67-97.

Weijer, J. van de (1999) *Language Input for Word Discovery*. Doctoral dissertation, University of Nijmegen.

Werker, J. (1991) The ontogeny of speech perception. In: I. Mattingly and M. Studdert-Kennedy (eds.) *Modularity and the Motor Theory of Speech Perception*. Mahwah, NJ: LEA, 91-116.

Werker, J. and R. Tees (1984a) Cross language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development* 7, 49-63.

Werker, J. and R. Tees (1984b) Phonemic and phonetic factors in adult cross-language speech perception. *Journal of the Acoustical Society of America* 75, 1866-1878.

Wiese, R. (1996) *The Phonology of German*. Oxford: Oxford University Press.