Phonological Fillers: Data from Norwegian

Hanne Gram Simonsen
University of Oslo

1. Introduction

Phonological fillers, i.e., phonological elements with no or uncertain semantic content, is a phenomenon found in many children's early production in a number of different languages (cf., e.g., Peters 1997). This paper addresses the question of how to interpret such fillers, in the light of data from Norwegian. The goal is to explore the possibilities of establishing criteria for identifying a filler as an approximation to a grammatical word, as a protogrammatical morpheme or placeholder, and/or as a more purely phonological filler fitting the salient prosodic patterns of the language. The paper also discusses the possible interaction of prosody with morphology and syntax in the acquisition process.

2. Data

The Norwegian data on phonological fillers are taken from a study of the phonological development of two Norwegian children, a boy (Tomas) and a girl (Nora) between 2 and 4 years of age.

(1) Two Norwegian children, Tomas and Nora (Simonsen 1990):

<table>
<thead>
<tr>
<th>Tomas</th>
<th>Nora</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>I</td>
</tr>
<tr>
<td>II</td>
<td>II</td>
</tr>
<tr>
<td>III</td>
<td>III</td>
</tr>
<tr>
<td>IV</td>
<td>IV</td>
</tr>
<tr>
<td>V</td>
<td>V</td>
</tr>
<tr>
<td>VI</td>
<td>VI</td>
</tr>
<tr>
<td>VII</td>
<td>VII</td>
</tr>
<tr>
<td>VIII</td>
<td>VIII</td>
</tr>
</tbody>
</table>

(2;0) (2;3)
(2;2) (2;4)
(2;5) (2;8)
(2;9) (2;11)
(3;0) (3;2)
(3;3) (3;5)
(3;10)
(4;1)

These two children approached the acquisition task with different strategies: Tomas used a typically analytic or word oriented strategy from the beginning of the recordings, while Nora started out with more of a Gestalt or tune oriented approach (in the sense of, e.g., Peters 1983, 1986, Peters & Menn 1993). Tomas seemed to have a word-like unit as his point of departure, while Nora started with a target which was larger than the word.

Nora's strategy is evident in her frequent use of phonological extensions or vowel fillers: the insertion of an unstressed vowel with no – or at least uncertain – semantic content before and between the interpretable words. This phenomenon was widespread at ages 2;3 and 2;4, but had more or less disappeared at the age of 2;8. Some examples are found in (2):
(2) Nora (2;3): Phonological extensions: vowel fillers

(a) [e\textsuperscript{1} den æ\textsuperscript{1} one]\quad \text{V} \quad \text{DEN} \quad \text{V} \quad \text{UNDER} \quad \text{?og} \quad \text{er} \quad (?and) \quad \text{that-one (is) \quad underneath’}

(b) [e\textsuperscript{1} den æ\textsuperscript{2} lie\textsuperscript{1} de:\textsuperscript{1}]
\quad \text{V} \quad \text{DEN} \quad \text{V} \quad \text{LIGGE} \quad \text{DER} \quad \text{?og} \quad \text{?kan/skal} \quad (?and) \quad \text{that-one (?can/shall) \quad lie \quad there’}

(c) [\textsuperscript{1}e\textsuperscript{n} sit\textsuperscript{1} fan\textsuperscript{1}]
\quad \text{HUN} \quad \text{V} \quad \text{SITTE} \quad \text{FANGET} \quad \text{?kan/vil} \quad \text{‘she (?can/wants to) \quad sit(on) \quad the-lap’}

(d) [\textsuperscript{1}h\textsuperscript{n} e\textsuperscript{1} so/ver]
\quad \text{HUN} \quad \text{V} \quad \text{SOVER} \quad \text{?} \quad \text{‘she \quad (?) \quad sleeps’}

(e) [\textsuperscript{2}ike\textsuperscript{1} den æ\textsuperscript{2} fal\textsuperscript{1} se\textsuperscript{1}n\textsuperscript{1}n]
\quad \text{IKKE} \quad \text{DEN} \quad \text{V} \quad \text{FALLE} \quad \text{V} \quad \text{SENGEN} \quad \text{?har/skal} \quad \text{?i/av} \quad \text{‘not \quad that-one (?has/shall) \quad fall (?in/off) \quad the-bed’}

(f) [æ\textsuperscript{1} layst po ë\textsuperscript{1} de\textsuperscript{1}]
\quad \text{V} \quad \text{V} \quad \text{LYST PÅ} \quad \text{V} \quad \text{DET} \quad \text{jeg} \quad \text{har} \quad \text{?} \quad \text{‘I \quad AUX \quad want to \quad (?) \quad that’}

(g) [æ\textsuperscript{2} vis\textsuperscript{1} tæj]
\quad \text{V} \quad \text{VISE} \quad \text{DEG} \quad \text{?jeg/skal...} \quad \text{‘(?I/shall ) \quad show \quad you’}

The small, raised numbers in the transcription indicate the Norwegian word tones or word accents. There are two distinctive word accents in Norwegian, Word Accent1 and Word Accent2 – each with a characteristic pitch pattern linked to the accented syllable. In East Norwegian, the dialect investigated here, they are realised as follows: in WA1, the accented syllable has a low pitch, while in WA2, it has a fall – a high-to-low pitch. Each word tone is linked to a tonal foot, which starts with an accented syllable followed by one or more unaccented syllables (as indicated underneath the transcription). The word accents, as well as the prosodic pattern of Norwegian will be discussed more thoroughly in section 3.2 below.

Nora's vowel insertions make the units of her utterances consist of more phonological material than clearly identifiable target words. To compare, we can look at some examples of Tomas' utterances at 2;0 – his earliest registered datapoint:
(3) Tomas (2;0): no phonological extensions/fillers

(a) [²tʰaː buːth ²ʉːliːtʰeː]  
   TA      BORT       JULETRE  
   take   away       Christmas-tree

(b) [¹beːhen ¹ɡjoːte]  
   BABYEN   GRÅTER  
   the-baby  cries

(c) [¹han ¹voːr³]  
   HAN      VÅT  
   he       (is)  wet

(d) [²nete ... ¹soːɛn ... ¹muːsa]  
   NEDI    SJØEN    MUSA  
   down-in  the-sea  the-mouse  
   (= ‘the mouse (is) down in the sea’)

Tomas has no phonological fillers – neither here nor later in his registered development. He rather seems to adjoin units of a target-word size (but not always having all the words, nor ordering them in such a way that would make the utterance grammatically correct from an adult point of view, as can be seen in (3c) and (3d).) The examples from Tomas’ speech are included to illustrate that phonological fillers are not a necessary feature of Norwegian acquisition. However, since fillers are our focus of investigation, we will concentrate on the data from Nora to see how they can be interpreted.

3. Possible interpretations

   The first question is how to interpret these fillers: To what extent can they be attributed semantic meaning as (grammatical) morphemes, to what extent can they be counted as proto-morphemes or place-holders, and to what extent are they mostly phonologically, i.e., prosodically motivated?

3.1 Grammatical interpretation

   Looking first for grammatical meaning, we seem to have a gradient here, as illustrated in (4): from morpheme – via protomorpheme – to no possible grammatical interpretation.

   (4) Degree of grammatical content:  
       morpheme > protomorpheme > non-grammatical fillers

To approach the question, we first need to take a closer look at the actual forms – both the forms of the fillers used by Nora, and of the adult targets into which they might possibly develop.

   The vowels used as fillers by Nora are listed in (5):

(5) Vowel fillers: Nora (2;3) and (2;4)  
   [æ e ø a i]
These are listed according to frequency: [æ ɛ] are clearly the most frequent (22 instances each), followed by [e æ] (12 and 6 instances, respectively), while [a ı] are only found in a few cases (3 and 2, respectively). Note that not all the vowels in the phonetic repertoire of the child are used as fillers. The filler vowels are all relatively front and unrounded, they are among the most unmarked vowels in the target language, and – possibly with the exception of [ı] – also the sounds that are used as hesitation sounds. In Nora's speech they can't be interpreted as hesitation sounds, though. Hesitation sounds are normally associated with a certain length and some stress, but Nora's fillers are all short and unstressed.

As for the probable adult targets – the set of unstressed grammatical words in the adult language – they are listed in (6), with their general distribution. This is not an exhaustive list: only the most frequent grammatical words are included (all among the 45 most frequent words according to the frequency dictionary Heggstad (1982)), and at the same time only those are included which are considered probable targets in Nora's speech. (“Auxiliaries” is put within citation marks, since the forms listed in this column are not necessarily always strictly auxiliaries, but they are always unstressed.)

(6) Adult targets

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>VIL ‘will’ /vɪl, vi/</td>
<td>ER ‘is, are’ /ær, æ/</td>
<td>DET, DEN ‘it’ /de, den/</td>
<td>ET, EN ‘a’(indef) /et, en/</td>
<td>I ‘in’ /i/</td>
<td>OØ ‘and’ /ʊ/</td>
</tr>
<tr>
<td>KAN ‘can’ /kan/</td>
<td>VAR ‘was, were’ /var, va/</td>
<td>HAN ‘he’ /han/</td>
<td>DET, DEN DE ‘the’ (def., only bef. adj.) /de, den, di/</td>
<td>TIL ‘to’ /til, ti/</td>
<td></td>
</tr>
<tr>
<td>SKAL ‘shall’ /skal,ska/</td>
<td>HAR ‘has, have’ /har, ha/</td>
<td>JEG ‘I’ /jæj, jæ/</td>
<td>AV ‘of’ /av, a/</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MÅ ‘must’ /mo/</td>
<td>etc.</td>
<td>etc.</td>
<td>etc.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Distribution

| V | + | + | + | – | – | + |
| N | + | + | – | + | + | + |
| Pro | + | + | – | – | – | + |
| A | + | + | – | + | + | + |
3.1.1 Morphemes (acceptable targets) or protomorphemes: criteria

As illustrated in (6), some of these monosyllabic grammatical words consist only of a single vowel, like the conjunction og /œ/ ‘and’, and the preposition i /i/ ‘in’. But in addition, word final /r/ has a weak realisation in East Norwegian and often disappears in connected speech – a tendency which becomes even stronger in children, who generally find /r/ difficult to produce. This way, the auxiliary er ‘is/are’ may get a single vowel pronunciation /æ/. If we also take into account that in rapid, connected speech both the sounds /j/ and /h/ may easily get a very weak realisation, both har ‘has/have’ and jeg ‘I’ may have single vowel realisations, too, as /a/ and /æ/, respectively.

This means that in many cases it is possible to give the vowel fillers a relatively straightforward interpretation as “real” grammatical morphemes or at least acceptable targets: the criteria then being that they match the adult target distributionally, and well enough phonologically, in the sense that they are close enough phonological approximations to be acceptable.

(7) Criteria for acceptable target morphemes:
Match the adult target distributionally; close phonological approximation

This is the case, for instance, with the fillers presented without a question mark in the examples in (2), illustrated again in (8):

(8) [æ] meaning er /æ(œ)/ ‘are’ in (2a)
[æ] meaning jeg /(j)æ(j)/ ‘I’ in (2f)
[a] meaning har /(h)a(:r)/ ‘have’ in (2f)

These examples illustrate what may be considered “close enough phonological approximations”: the vowels are target-like, and the missing consonants are only those that are hard for the child to pronounce, and that furthermore may often be omitted in rapid speech.

However, often the match is not as good as this – either, the phonological match is not good enough (like the first vowel filler in (2a), where the possible interpretation as a conjunction ‘and’ would need a rounded /œ/), or the distributional match is lacking (like the third vowel filler in (2f) where no evident interpretation can be given).

3.1.2 Protomorphemes or non-grammatical fillers: criteria

If the vowel filler is a distributional match, but not a phonological match, we have a good case for a protomorpheme. But under a protomorpheme interpretation, we may make at least two slightly different assumptions:

1. The child utters an underspecified grammatical word – implying that there is actually a word-like grammatical target there for the child, but it is phonologically underspecified.
2. The child is not actually focussing on a single word, but rather filling out a larger syntactic structure – a pattern or schema with grammatically defined slots, one of which is filled by the vowel.
These two different interpretations correspond to different levels of schematicity or specificity in the sense of, e.g., Langacker (1987). (Possibly, there may even be more than just these two levels of specificity/schematicity, but these two at least seem identifiable here.) So here again, there is a gradient. This means that we should replace (4) by (9), where the protomorpheme may be interpreted more specifically as, e.g., a protomodal, a protopronoun, or a protoauxiliary, or it may be interpreted more generally, as a placeholder of a more underspecified or schematic nature.

(9) **Degree of grammatical content – revised**

<table>
<thead>
<tr>
<th>morpheme</th>
<th>protomorpheme</th>
<th>non-grammatical filler (placeholder)</th>
</tr>
</thead>
<tbody>
<tr>
<td>protomodal</td>
<td>protopronoun</td>
<td>protomorpheme</td>
</tr>
</tbody>
</table>

To decide between these, more specific criteria are needed. As these must often be language-specific, we shall not go into details here, but look at some examples to illustrate the point:

(10)(a) [hun ε ²laːt opo ¹den] HUN V LAGT OPPÅ DEN she ?AUX laid(ptc) on-top that-one

‘She ?is laid on top (of) that/ she ?has laid (it) on top (of) that’

(b) [ɛ ¹leːse ¹buːk] V LESER BOK ?PRO read book

‘I /you/we read (the/a) book’

(c) [ɛ ¹den ε ²ligɛ ¹daː:] V DEN V LIGGE DER ? that-one ?MOD lie (inf) there

‘?and that one ?can/shall/may lie there’

(d) [æ ²viːs tæj] V VISE DEG ?MORPH show you

‘?I /?can/shall show you’

In the position in front of a verb in Norwegian, modals, auxiliaries, pronouns, and conjunctions may all occur – as indicated in (6) above. But there are restrictions regarding the form of the verb:

1. If the verb is a **participle**, a **protoauxiliary** reading of *er ‘is’ or har ‘has’* is very probable – cf (10a).

2. If the verb is **finite**, i.e., a present or past tense form, the filler cannot be a modal or an auxiliary. So if it is possible to identify the verb as finite (which is not always easy in a child’s production) – we may adopt a specifically **protopronoun** interpretation. This is the case in (10b), where the verb may be interpreted as a finite, present tense form in spite of its missing present tense /r/ suffix, because of the word accent pattern: the verb *lese ‘read’* is among the verbs which has WA2 in the infinitive, but WA1 in the present tense.

3. If the verb is an **infinitive**, both modals and pronouns are possible. In that case, the case is stronger for a **protomodal** if a pronoun is clearly present (cf (10c)). Otherwise, we must remain with the less specific, more schematic
placeholder or **protomorpheme** interpretation, like in (10d), where both a pronoun (e.g., ‘I’,) and a modal (e.g., ‘can’ or ‘shall’) are possible and probable readings.

This gives us the criteria for identifying protomorphemes listed in (11):

(11) **Criteria for protomorphemes:**

- Match the adult target only distributionally, not phonologically.
- Degree of specificity/schematicity depends on how closely the distribution can be determined.

### 3.2 Phonological/prosodic interpretation

Using the criteria for interpreting fillers as acceptable target morphemes in (7) and as different kinds of protomorphemes in (11) above, a good portion of the vowel fillers in Nora’s utterances may be accounted for, and given some degree of grammatical meaning – as illustrated in (12). (Note that this table includes only the fillers in Nora's utterances, so that the column “Acceptable target morpheme” does not include all the grammatical morphemes in her speech, but only those of her fillers which can be given an interpretation as an acceptable attempt at a specific grammatical morpheme.)

#### (12) Filler classification and development, multisyllabic utterances

<table>
<thead>
<tr>
<th></th>
<th>Acceptable target morpheme</th>
<th>Proto-aux/mod/pron/art/conj/prep</th>
<th>Proto-morpheme</th>
<th>Non-grammatical</th>
<th>Total number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nora 2;3</td>
<td>27</td>
<td>16</td>
<td>20</td>
<td>6</td>
<td>69</td>
</tr>
<tr>
<td>Nora 2;4</td>
<td>6</td>
<td>10</td>
<td>10</td>
<td>3</td>
<td>29</td>
</tr>
<tr>
<td>Nora 2;8</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>5</td>
</tr>
</tbody>
</table>

There is a clear decrease in the total number of fillers used, mainly due to a development towards a more distinct pronunciation, resulting in a closer phonological match of the grammatical morphemes so that they do no longer have the character of fillers. (Unfortunately, there was a gap of more than 3 months between the second and the third session, so that it is not possible to trace the actual transition more closely).

But as we have seen, to some extent the grammatical meaning given here is of a very general nature, and in addition, not all the vowels can be given any grammatical meaning at all. (Those are the ones marked without a gloss, having only question marks in the examples in (2), f. ex. the [e] in (2d) and the third filler [ɛ] in (2f).) In these cases, a more purely prosodic interpretation is called for.

According to this interpretation, the vowels are inserted to fit a certain rhythmic pattern which is salient to the child. A quick glance back at the examples in (2) could suggest one, relatively simple rhythmic principle: namely that the vowels are inserted to avoid a “stress clash”, as stated informally in (13):
Prosodic principle I: Stress clash avoidance

Avoid two stressed syllables adjacent to each other

Stress clash avoidance has been presented as a general, possibly universal phonological principle (Liberman & Prince 1977, Hayes 1984). On a phonetic basis, it has been difficult to establish empirical evidence for such a universal principle — however, a general cognitive ability to give rhythmic interpretations to auditory material may form the basis for perceptions of rhythm in speech (cf. the discussion in Laver 1994: 523f). Whether the languages are predominantly stress-based (like most Germanic languages) or predominantly syllable-based (like most Romance languages), children may form hypotheses of rhythmic patterns on the basis of their perception of the input they receive. In a stress-based language like Norwegian alternations of stressed and unstressed syllables may well present a salient pattern to the child acquiring the language.

Indeed, there seems to be a strong tendency in Nora's utterances to avoid stress clashes. We do find a few examples of them, like in (14):

(14) Stress clashes in Nora I

(a) [iin ɪˈdaːt]  
   - -  
   INN DER  
   in there

(b) [ɪˈdæn ɪˈveːt]  
   - - -  
   DEN HAR VI  
   that-one have  
   we (= ‘that one, we have’)

But they are few: out of 176 multi-syllabic utterances in the recording of Nora at age 2;3, only 6 have a stress clash. And as shown in the examples in (2), very often a vowel is inserted where there would otherwise have been a stress clash.

But in addition, as mentioned in section 2 above, in Norwegian not only the stress pattern, but the pitch pattern is salient, enhancing the perception of rhythm. Norwegian (like Swedish) is a stress-based language with a word tone or word accent system which works in interaction with the intonation system (cf. Gårding & Stenberg 1990, Fretheim 1992). The word accent carries mainly lexical information, while the intonation primarily carries syntactic and pragmatic information.

The word accent oppositions are associated with the accented syllables: Every accented syllable carries the pitch pattern of one of two possible word accents, WA1 (a low tone) or WA2 (a falling tone). To have an explicit contrast between the two accents, the accented syllable must be followed by at least one unaccented one, a fact that ties up well with the stress clash avoidance principle. (A monosyllabic foot is always WA1). Since the difference between the two word accents is not directly relevant here, we will not go further into details about the word accents, but only note that the word accent is tied to the basic prosodic unit at utterance level, the tonal foot, starting with an accented syllable and including all syllables up to the next accented syllable (Cruttenden 1986, Fretheim 1991, 1992). The pitch pattern of the unaccented syllable or syllables following the accented ones may vary, and belongs to the sentence intonation domain.

To sum up, as illustrated in (15): A prosodically well-formed utterance in Norwegian is built up of intonational phrases, each consisting of one or more
Tonal Feet (TF). Each Tonal Foot starts with an accented syllable, optionally followed by one or more unaccented ones. Optionally, the first Tonal Foot of an utterance may be preceded by an Onset (O) consisting of one or a few unaccented syllables. (Since Nora's utterances at this point each seem to contain only one Intonational Phrase, we will not discuss the difference between these two levels of the prosodic hierarchy here.)

(15) Prosodic hierarchy for Norwegian:

\[
\text{Intonational Utterance} \\
\text{|} \\
\text{Intonational Phrase} \\
\text{|} \\
\text{Onset} \\
\text{|} \\
\text{(TF) ...} \\
\text{Tonal Foot} \\
\text{|} \\
\sigma ... \\
\text{σ} ... \\
\text{σ} ... \\
\text{σ} \\
\text{[–acc]} \\
\text{[+acc]} \\
\text{[–acc]}
\]

This model offers several possibilities for prosodically well-formed utterances in Norwegian. However, the most frequent pattern in East Norwegian is a pattern where most Tonal Feet are filled out with unaccented syllables. A count in the adult speech to Nora at 2;3, for instance, shows that among the 390 multisyllabic Intonational Phrases found there, 84% have full Tonal Feet, (i.e., feet having an accented syllable followed by at least one unaccented one) at least initially and medially. Thus, one may hypothesise that such a pattern will be salient for the child – an entrenched prosodic schema to put it in cognitive linguistic terms.

And indeed Nora's multisyllabic utterances seem to follow this prosodic schema, filling out the Tonal Feet. As illustrated in (16) below, this tendency is very strong: more than 95% of her multisyllabic utterances consist of full, multisyllabic Tonal Feet. Her vowel insertions often fill out an otherwise monosyllabic Tonal Foot.

In addition, she also shows a somewhat weaker tendency to start her utterances with an unaccented Onset consisting of one or more syllables. Approximately 60% of her multisyllabic utterances have such an Onset, and among these again, more than 60% consist of a single vowel filler. To compare, in the adult speech sample to Nora at 2;3, 71% of the utterances have an Onset of this kind – another relatively salient pattern, but evidently less entrenched for Nora.

(16) Multisyllabic utterances:

<table>
<thead>
<tr>
<th></th>
<th>Adult speech to Nora (2;3)</th>
<th>Nora (2;3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full Tonal Feet</td>
<td>84 %</td>
<td>95 %</td>
</tr>
<tr>
<td>Onset syllable(s)</td>
<td>71 %</td>
<td>60 %</td>
</tr>
</tbody>
</table>

These tendencies or possible schemas in Nora's speech are summarised in (17) – (where Schema I also incorporates the stress clash avoidance principle):
(17) **Prosodic schemas for multisyllabic utterances:**

I Fill out Tonal Foot to make it multisyllabic
II Start utterance with Onset syllable(s)

4. Grammar and prosody: Interaction or exclusion?

I started by trying to discover strict criteria for identifying the different levels of grammatical interpretation of phonological fillers, illustrated by the vowel fillers in the speech of the Norwegian girl Nora. However, in doing this I do not imply that the prosodic interpretation of the fillers is restricted only to those cases where no grammatical interpretation can be given. Rather, as suggested by, e.g., Peters & Menn (1993), the prosodic pattern or schema is there the whole time as a baseline – a pattern to tune in on, for example for a tune oriented child like Nora.

But then she must learn to interpret this prosodic pattern: In a Norwegian tonal foot, most of the lexical information (normally that of one lexical item) is carried by the accented syllable, while the grammatical (morphological and syntactic) information is carried mostly by the unaccented syllables (as derivational or inflectional morphemes, or as separate grammatical words).

Once the child realises to some extent or at some level what word class the lexical item belongs to, she can start building up information of the morphosyntactic meaning carried by the surrounding unaccented slots. Without having to know the exact phonetic details of the grammatical units, the child may use the prosodic pattern as a scaffolding to the full-fledged grammatical system, which may be built up gradually through use. And as indicated above it seems that Norwegian is a language where prosody interacts with morphology and syntax in a way which makes such a filler strategy useful.

However, this is not the only possible strategy – as exemplified by Thomas, the other child in my study, who used virtually no fillers at all! But that is another story...

Endnote

* I wish to thank the participants at the IASCL symposium on fillers for their valuable comments and suggestions, in particular Conxita Lleó, Susana López Ornat, and Ann Peters. In addition, Rolf Theil Endresen, Thorstein Fretheim, and Inger Moen have provided many useful discussions and comments on earlier versions of this paper.
References


