Filling the Word Gap
An Experimental Typology of Infant Utterances Between Babbling and Speech

Tim Schmitz-Reinthal

Introduction
This article evolved from a project initially dealing with the articulatory features of first language acquisition. While working with that project the problem arose of how to separate utterances which are worth examining and those which are not. I then changed the emphasis of my studies and began to research the field of utterance classification in the early stages of L1-acquisition. As I discovered, there was a lot of pioneering work to do, because as soon as there are the first words available from a first language learner no-one seemed any longer to be interested in the previous stages of speech which, however, were still in use.

Background (Why words?)

“...In studies of early child language the word has been tacitly accepted as a basic unit for analysis.” (Vihman & McCune 1989:1)

Words, even if they are not easy to define in linguistic terms, have often been used as the basic object in language acquisition research. They have been seen in contrast to babbling. An invisible line has been drawn between ‘speech’ and ‘prespeech’ and that implies a step from speechless babbling children to those which are able to handle words. Research has concentrated on that part of utterances that is close to the words used by adults, which have been used as a standard. Jakobson 1942 and many researchers after him have described the development of a mother tongue as discrete steps and not as a continuum.

While working on the project mentioned above, dealing with the acquisition of phonemic features together with the acquisition of semantic units (words), some questions began to arise: Why words? How to define them? Are there only words and non-words?
Why words? – Words are units which are located on the surface of language, they are not as 'constructed' as e.g. morphemes or other linguistically defined units and they are transparent to the layperson. Words are structures which are easily recognizable by children – in particular since parents often try to teach their children just words. It is therefore no wonder that researchers devoted most of their attention to words.

What is a word? – That is a good question. The ubiquitous 'between-two-blanks'-definition does not seem to function in the field of first language acquisition research. Labov & Labov 1978:821 mention three factors to determine a child's word:

- Clearness of articulation and accuracy regarding the placing of vowels,
- Distance (in time) from other sounds or sound chains,
- Referents which ensure identification.

But it is not the aim of this article to define the meaning of 'word', and because of its vagueness and ambiguity this term has not been used to a great extent. When it does occur then only with the meaning 'utterance with certain semantic content'. Where there is any need to be more precise, terms like 'adult's word' or 'model' were used for utterances made by adults or 'target' or 'empty word' for certain child utterances.

Can words be defined as discrete units which either do or do not exist? – That is the aim of this article, to show that there is more than just words and babbling. Even if it seems that the space between these two extremes consists only of more clearly defined intervals, the problem of type-token-coordination and that of homonymy/homophony shows that the space between babbling and speech must be seen as a continuum.

The goal of this classification is therefore not to classify words already found in a subject's speech acts nor is it to stand as a manual to classify utterances. It should instead illuminate the continuum mentioned above where there is a gap in traditional research.

Jakobson's thesis of discontinuity does not seem to show the reality of language but the interests of science (see Jakobson 1942).

The Corpus
The material consists of about 25 recording hours of two girls between the ages of 6 and 18 months, both of them are acquiring Swedish as a first language, one of them also German. Any utterances which could be defined as 'speech' or 'prespeech' were transcribed, but not vegetative utterances (weeping, laughing, crying etc.).

Malin (the bilingual child) was recorded over a period of eight months, Sofia, the other subject, over a period of six months. (Malin was recorded from the age of 9 to 18 months, Sofia from 6 to 12 months.) The interval between recording sessions was about eight days (±4) for Malin and two weeks for Sofia. The medium length of sessions was 30 minutes for the first six sessions with Malin and the two last sessions with both subjects. For any other session the duration was about 45 minutes. Together there were 24 recordings with Malin and 8 with Sofia.

The recording technique was the same for both subjects: A wireless microphone granting greatest possible freedom of action. The recording device was a Sony cassette-recorder. No video-recordings were made.

During the sessions the child's mother and the interviewer were usually present. Both were partially dealing with the child and partially in the background. Sometimes there were more persons present. The children played with all things at hand from toys to recording devices.

Recordings were taken inside in the child's natural living environment, with the aim that there should be as few disturbances as possible, and new and unfamiliar situations should not occur. The recordings were transcribed and transferred into digitally readable form.

Type, token and category
A classification of tokens is based on synchronic evaluation while types are diachronic in their nature. During the work on transcription it was a goal to develop rules general enough to satisfy the demand of an instant evaluation. It soon became evident that this was not possible. Classification was mostly done on criteria which were based on previous experience with similar situations and/or the phonetic form of the token. Each token seemed so to be tied to a type or a certain category. It was often difficult to decide how they belonged together, because this classification depended on intuition. It became clear that putative types were not homogenous at all. Tokens which seemed to belong to a certain target often only had a similar phonetic form.

In this article the utterance groups defined below are called category (I to IV) while type is just a subcategory in the target category consisting of target tokens.

Evaluation of utterances
Vihman & McCune 1989 designed a series of criteria to decide if an utterance can be defined as meaningful. These were very practical. The
utterances in the Lund corpus were first evaluated intuitively, but after a thorough analysis the following criteria crystallized to define types:

- context (speech and non-speech)
- phonetic similarities with target
- parents identify new words
- imitation
- semantic expansion of existing words
- intensity of use

Vihman & McCune use the first three features and the fourth in some special cases when imitation cannot be seen as type imitation as it is defined below. But there are even other divergences between their method and the one used here:

Deictics. Vihman and McCune say, cannot be identified contextually. This however was done in the Lund corpus. Words like där 'there' (Swed.), titta 'look' (Swed.) and da 'there' (Germ.) have high frequency rates and were always identified by contextual matters if the situation permits seeing them as part of a deictical action. If an utterance did not fit into a contextual frame, it was not identified as belonging to the possible target type.

<table>
<thead>
<tr>
<th>Malin:</th>
<th>['ti:ta'] (look)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mother:</td>
<td>'titta?' +++ 'guck mal' heißt das (look? +++ it is 'guck mal')</td>
</tr>
<tr>
<td>M:</td>
<td>['ke 'wui']</td>
</tr>
<tr>
<td>Interv' er:</td>
<td>ja + das war schon ganz gut (yes + that was really ok)</td>
</tr>
</tbody>
</table>

(Malin – session 6; 18/11/91; tape no. 3b, 131-132)

Phonetic similarities were split up into three identification classes – complex match, exact match and prosody – even here following Vihman & McCune. They say in regard to complex match that this is valid for tokens matching with more than two segments, which, however, is rather arbitrary. This criterion was used here only together with others, so I did not set any specific restrictions. As shown in (1) there is almost always a coincidence of features. Here it is match in prosody and initial and imitation.

Link between evaluated tokens cannot be neglected. The activity level of similar tokens and their embedding in special contexts suggests that the tokens definition can be used in certain situations.

The expansion of the semantics of already existing types can lead to a sort of type that does not match with the semantics of the corresponding

<table>
<thead>
<tr>
<th>Table 1. Categories of the speech acquisition continuum</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \text{Ia babbling} )</td>
</tr>
<tr>
<td>meaning</td>
</tr>
<tr>
<td>-</td>
</tr>
<tr>
<td>meaning</td>
</tr>
<tr>
<td>0</td>
</tr>
<tr>
<td>meaning</td>
</tr>
<tr>
<td>-</td>
</tr>
<tr>
<td>meaning</td>
</tr>
<tr>
<td>-</td>
</tr>
<tr>
<td>meaning</td>
</tr>
<tr>
<td>-</td>
</tr>
<tr>
<td>meaning</td>
</tr>
<tr>
<td>-</td>
</tr>
</tbody>
</table>

IVa, b are targets, IVc, d are phatic imitations.

meaning: (+) has a concrete or situational meaning, (-) no meaning;
freedom: (+) can be defined by situational context, (-) not evaluable without context;
phonetic: (+) matches a model in discrete phon. units, (-) no model matching;
prosody: (+) matches a model in prosody, (-) no model matching;
activity: (+) tokens are frequent in infants speech, (-) tokens is new or unusual;
(0 - anywhere): plays no role for evaluation or both (+) and (-) appear.

Target. However, as long as the phonetic match does not change it will be classified as the same type as before (see homophony etc.).

Classification of utterances

The identification and classification of utterances is above all based on semantic and phonetic (articulatory) items. That means, one has to decide if an utterance can be seen as a bearer of meaning, something that often only depends on the local situational context. Of some importance is also the relation between the phonetics of the child’s utterance and the one of the target. If we look at semantics we can assume that children in that age of first-language acquisition very seldom refer to things or concepts other than those defined by the situational context (here and now).

If one looks at phonology and semantics as distinctive features, the following table of four groups can be established:

- **category I**: babbling (- meaning/ - articulatory similarities with a model),
- **category II**: jargon (- meaning + articulatory similarities),
- **category III**: protowords (+ meaning / - articulatory similarities),
- **category IV**: target words (+ meaning / + articulatory similarities).

These categories can be divided into the subgroups defined in table 1.
‘Meaningless’ utterances

Babbling (category I)

The term babbling here means utterances without any transparent connection to the child’s situation and without any concrete or deictic meaning, which are not used in a playful intention, which are not imitated and have no or almost no similarities to adults’ words.

It is not as easy as it seems, however, to clearly separate babbling from utterances carrying meaning or speech acts. Tokens are often found which in an articulatory sense are similar to those belonging to a type in the child’s lexicon. They are, however, not very meaningful in the situational context and differing in this from ‘real’ target words. They were called here empty words, showing that they seem to have a distinct meaning, because phonetic form suggests this, although they are semantically empty.

Jargon (category IIa)

(2) Malin: is playing with various toys, shuffling them back and forward again

Interv’er: und jetzt gehst du bodenwischen? (and now you are going to sweep the floor?)

(Malin – session 9, 91-12-19, tape no. 5a, 076)

Jargon emerges from a child trying to mimic an adult’s utterance that is not intended to be said to the child. Adults often think that the child expresses something special that they do not understand, especially when the child’s utterance prosodically resembles fluent speech.¹

Jargon can be a direct imitation or can be produced during other activity like in (2). The child does not seem to be aware of its parents or other persons recognizing this wrongly as an intended speech act.

Sometimes jargon develops from situations when a child wants to take part in a dialogue or an action going on without any plan to integrate the child, e.g. phone calls.

(3) Malin: is playing with his lips

Interviewer and Malin are looking in a picture book, mother is speaking on the phone, groans, is talking on the phone

(Malin – session 9, 91-12-19, tape 5a, 199-202)

¹ Malin’s German aunt was told the story about this: she interpreted such a jargon expression as if the child had said something in Swedish.

‘Meaningful’ utterances

Target words (category IV a, b)

Expressions that have adults’ words as a model are here referred to as targets. They should therefore match the model in its phonology and semantics – i.e. they should at least once have the same meaning as the model.

There may, however, be cases with ‘irregularly’ extended semantics. The target da ‘there’ (Germ.) has for example here in addition to its deictic character more requestive meanings (see Carter 1979:78) as ‘please’ and ‘thank you’.

The target category features two major subcategories, as there are ‘real words’ and ‘baby talk’. The model for ‘reals’ is a typical everyday adult’s expression as e.g. da ‘there’, hej ‘hi’ or Ball ‘ball’. These can be further separated into words referring to concrete things, actions, etc., and deictic expressions or those which refer to situational concepts. Carter 1979 is one
author who gives more subclasses of types, but the classification of types as
defined above is not a goal of this work.

(5) Mother: na + endlich ist sie fertig
    (well + she is ready at last)
Interv’er: na + endlich ist sie fertig
    (well + she is ready at last)
Malin: [əja]
Mother: ajajaj + det var pappa + endlich ist sie fertig
    (aye aye aye + it was daddy + she is ready at last)
(Malin – session 3, 91-10-29, tape 2a, 550)

(6) Malin: points at the ball
    [paː bəbə]
Interv’er: ja das ist der Ball
    (yes that is the ball)
(Malin – session 6, 91-11-18, tape 3b, 080)

To distinguish these from type imitations – forming a class of their own –
it is of importance to pay attention to the fact that there may be imitated
tokens of a target but the concept of the word must be independent.

Besides the so called ‘reals’ there is ‘baby talk’, emerging from mother­
ese utterances of parents or other adults. These utterances are of temporary
nature both in adults and in children and are likely to disappear in a later
stage. In spite of their temporary nature, baby talk expressions are clearly
defined as targets, as the child recognizes motherese as the model and goes
through the same design process as with ‘real words’.

(7) Father: de iñ aja
    (they are ‘no-no’)
Interv’er: ist alles verboten + ne + alles ist verboten
    (is everything forbidden + isn’t it +
everything is forbidden)
Malin: [əja]
(Malin – session 5, 91-11-13, tape 3a, 451-455)

(8) Malin: discovers microphone and wants
to play with it
    machen wir’s mal weiter nach hinten +
    damit du’s nicht so gut abkriegst
    (let’s put it farther back + so you cannot
    pull it off)
M: [ŋa]
I: ja
    (yes)
(Malin – session 10, 92-01-06, tape 5b, 076)

Type imitations (categories IIc, Vic, d)
The term imitation may be a bit misleading, because even targets deserve
this description – they often emerge out of an imitation of a model. It must,
however, be differentiated between ‘echoic imitations’ of certain target
types and more extensive imitations which actually cover the range of a
whole type. The latter version makes up the category called type imitation.
There are two different possibilities of type imitations:
   – The child imitates utterances from its surroundings unusual to language,
     e.g. the mimicking of a motor sound. This may even be only playful
     (see game sounds). A very interesting version of imitation appeared when
     Malin began to imitate the family’s cat. This is illustrated by example (9).
   – The child’s lexicon contains a type with any tokens realized only as
     imitations in response to an approach.

Protowords (category III)
Ferguson 1976 introduced the term protoword as a type with infant-defined
(independent) phonology. The concepts of protowords do not have to be
unique to normal language.

There are no examples for plain protowords in the Lund corpus except
for the ‘cat moan imitation’ mentioned above:

(9) Malin: [ha ha’ha]
    Mother: ja + das war die katze was hat die katze gemacht
    (yes + that was the cat what has the cat done)
Mal: [hehe ‘a]
    cat meows, pause
    happily

Special problems
Homophony, polysemy, synonymy
Lise Menn 1978:10 expresses the view that there is – analogous to the step
theory of babbling vs. words – a clearly defined step between meaningful
and meaningless utterances: “... if the meaning does not recur when sound­
patterns recur, we have babble, and if sound-patterns do not recur when a
recurrent meaning is apparently not intended, we do not have linguistic ex­
pression of that meaning....” This simplifies the problem of homophony
etc. Due to the opacity of the infant’s inner representation of articulator}'
features, meaning, and the link between them, it is not clear whether they
are of the same sort as in adults or, if they are, whether the child is able to
express them in this way. A child may have two synonymous expressions
where adults have only one (see also Two L1), or s/he may have a phonetically stable type which can expand its semantics.

Infants’ utterances are almost always ambiguous. Their interpretation depends strongly on the interpreting persons and even intrapersonal interpretations can contradict each other. On the other hand already identified types can cover a larger or smaller part of the semantic space than the adult’s word in question.

Homophony, etc., can cause irritation in more than one way. One of them is polysemy of a target type. Malin often used period utterances belonging to the German model da ‘there’ (realized as e.g.[da], [ta], [tae]), phonetically and semantically during the investigation. After a while more and more tokens showed up which could be classified as phonologically identical, but they did not match semantically. Since they were often used in situations where something was handed over, new meanings such as ‘please’, ‘here you are’, and ‘thank you’ arose.

Another one of the irritating cases is homophony of tokens belonging to different categories. It often happens that a token turns out to be better classified in a totally different category. Typical for that are the so called empty words – utterances which are amazingly like targets until it becomes clear that they do not fit at all into the situation they are uttered in.

Finally one must mention a splitting of a meaning into two phonological realizations and their eventual reunification at a later stage. The question is how to treat the tokens concerned. Do they belong to one, two or even three types (two before and a third after reunification)? Should one of the types be erased retroactively? The problem has not yet showed up in the current investigation, but it may become relevant with respect to da ‘there’ (Germ.) and där ‘there’ (Swed.).

Two L1
If there exists one model in both languages or if there are two satisfactorily similar words in both lexicons which can act as a model, there will be complications in relating the child’s output to the specific language. It is not clear whether there is a representation of two languages in the infant’s mind at all at this stage. In this case it was assumed that a child mainly interacts in the language it is spoken to, but there may be some cross-language interactions as in (1).

In cases where possible targets exist in both languages, with very little difference in phonetic form, the classification appears to be rather arbitrary. So [da] may refer either to da or to där. In the present case of Malin it was assumed that two types exist, because of her consistent realization of da as [da].

Phatic communication with words
Neither children nor adults use language only for communication via a defined lexical meaning. Language is even used to express moods, demand attention etc., anything which could be read ‘between the lines’. Phatic communication is communication for its own purpose. Many utterances made by infants can be defined as phatic. They can theoretically belong to any of the categories named above.

It is typical that a child is in a situation, where it must demand attention. This can be done by any articulation, bearing semantic information or not. It happens very often, however, that phatic use is emphasized by words fitting into the situation:

(10) Mother: wants to have attention
    Malin: $[^{\text{ja}}]$ (hi/yes?)
    Mother: ja + genau (yes + exactly)

(Malin – session 11, 92-01-13, tape 6a, 503)

Summary
As proposed here the gap between babbling and those expressions usually called first words can be described more or less as a variety of utterances approaching adult’s language. The classification used in this study is a first attempt to make this space transparent, so the categories defined can be seen as proposals. As this is an experimental approach there is a lot of work to do to establish a sufficient theory of utterances filling the ‘word gap’.

Acknowledgements
I would like to thank first of all my subjects Sofia and Malin and their parents for their patience, help and understanding. I would also like to thank the staff from Lund University Dept. of Linguistics, especially Gisela Håkansson and Bengt Sigurd for tutoring me during my research project and John Peterson from Kiel who corrected my manuscript. Finally I have to thank the Swedish Institute for financing my research during the last year.
References

The Problems of Multi-word Prepositions and Subjunctions
Bengt Sigurd

Introduction
Multi-word prepositions such as Swedish i mitten av, mitt i, English in the middle of, Swedish på grund av, English because of, Swedish i överensstämmelse med, English in accordance with, according to, Swedish vad beträffar, i fråga om, English as to, are quite common in written and spoken texts, but they are rarely mentioned when the category of prepositions is discussed in traditional or modern linguistics. One reason for this deficiency is probably that they consist of several words and such items cause problems for lexicographers and grammarians. They don’t come to mind as easily and tend to be disregarded.

The purpose of this paper is to draw attention to these multi-word prepositions and suggest ways of treating them in formal grammar and automatic translation. A number of multi-word subjunctions will also be treated. Many of the subjunctions can be said to be formally related to prepositions by the addition of att in Swedish, the fact that in English, e.g. Swedish tack vare (preposition), tack vare att (subjunction), English thanks to, thanks to the fact that, Swedish bortsett från (preposition), bortsett från att (subjunction), English disregarding, disregarding the fact that, Swedish trots, till trots av (preposition), trots att, till trots av att, English in spite of, despite the fact that.

The expressions illustrated could be analyzed and understood as combinations of the individual words already in the lexicon. A look at the examples mentioned shows that this is possible in some cases, but a clumsy if not impossible approach in others: there is no common free noun spite from which in spite of can be understood and it would be most difficult to derive the meaning of as to from as and to.

This paper suggests that a number of recurring word sequences of the types illustrated are best treated as multi-word prepositions or subjunctions and included as separate entries (lexemes) in the lexicon.