Research Repository UCD

Title	Identifying formulas in first language acquisition
Authors(s)	Hickey, Tina
Publication date	1993-02
Publication information	Hickey, Tina. "Identifying Formulas in First Language Acquisition" 20, no. 1 (February, 1993).
Publisher	Cambridge University Press
Item record/more information	http://hdl.handle.net/10197/4167
Publisher's version (DOI)	10.1017/S0305000900009107

Downloaded 2023-04-07T10:52:29Z

The UCD community has made this article openly available. Please share how this access benefits you. Your story matters! (@ucd_oa)

© Some rights reserved. For more information

Identifying formulas in first language acquisition*

TINA HICKEY

Linguistics Institute of Ireland

(Received 6 August 1990. Revised 3 December 1991)

ABSTRACT

With the increase in interest in formulas, or apparently non-productive utterances in children's speech, a range of definitions has emerged and sometimes conflicting criteria have been proposed for their identification. These definitions of formulas are compared, and the criteria of Brown (1973), Wong Fillmore (1976), Peters (1983) and Plunkett (1990) for the recognition of formulas are reviewed. A preference rule system is proposed, which distinguishes necessary, typical and graded conditions for the recognition of formulas. Using these conditions, some of the formulas found in the data of one child acquiring Irish between 1;4 and 2;1 are examined. Issues such as length of units, frequency of occurrence and appropriateness of use are discussed. The methods developed in this study could be used to assess the importance of formulas in the language acquisition of other children.

INTRODUCTION

While the existence of certain memorized units in children's speech has long been acknowledged (e.g. Brown, 1973, and Bloom, Lightbown & Hood, 1975), the emphasis in child language research has generally been on productive language. Unanalysed chunks of language whose elements are not productive have been called 'formulas'; a formula functions as a unit for the child. Some of the theories regarding formulas in acquisition will be reviewed briefly here. Then the crucial issue of reliable identification of formulas will be considered, principally discussing the criteria of Brown (1973), Wong Fillmore (1976), Peters (1983) and Plunkett (1990). Finally, some of the contentious issues in the recognition of formulas will be discussed in the light

[*] Earlier versions of this article were presented at the Ninth World Congress of the International Association of Applied Linguistics (AILA) 1990 and the Fifth International Congress for the Study of Child Language 1990. I am grateful to Michael Garman, Bernard Spolsky and other conference participants for their comments. Address for correspondence: Tina Hickey, Institiúid Teangeolaíochta Éireann/Linguistics Institute of Ireland, 31 Fitzwilliam Place, Dublin 2, Ireland.

of examples from a study of the acquisition of Irish as first language by one child from the age of 1;4 to 2;1.

Jespersen (1924) recognized the role of formulas in adult speech, which he defined as fixed expressions in which neither stress, rhythm nor words can be altered, and which are acquired whole. Bolinger (1976) argued that, in adults, formulas become established because they express emotions or situations that are frequently experienced, acting rather like the old printing practice of keeping together as a block (or cliché) those phrases which are used often. He claimed that formulas allow greater fluency than would be possible if every phrase had to be newly minted.

Wong Fillmore (1976) adopted Jespersen's distinction between formulas and productive speech in her analysis of the development of English as a second language in young Spanish-speaking children. However, she broadened the definition of a formula, stating that it referred to 'language which functions wholly or partly as unanalysed, fixed or automatic units for the speaker' (1976: 295). Thus her definition allowed for the notion of a formulaic frame, a formula which has been partly analysed, so that there is some substitutability in a grammatical slot within the formulaic construction. Such a formulaic frame may eventually either become fully analysed into its freed components, or remain only partly analysed. She claimed that her subjects used their formulas as data on which segmentation and analysis were later carried out, since this language which they had already 'caught' was most accessible for analysis.

Wong Fillmore's study of formulas was based on data from children acquiring English as a second language. Other studies have investigated the role of formulas or formulaic expressions in first language acquisition. Clark (1974) and Peters (1983) have proposed that formulas play a significant role in the acquisition of some children. Several authors (e.g. Peters, 1983; Bretherton, McNew, Snyder & Bates, 1983; Bates, Bretherton & Snyder, 1988; Hampson & Nelson, 1990) have noted significant individual differences in children's use of formulas. This has been linked to the differences in style found by Nelson (1973) and Bloom *et al.* (1975). These were later termed 'referential/nominal/bottom-up' and 'expressive/pronominal/top-down' by Bretherton *et al.* (1983). Nelson (1981) observed that expressive/pronominal children used more personal-social words and formulas both in the one-word stage and in the early stages of grammar than did the referential/nominal children, who favoured concrete nouns.

The role of formulas in language acquisition has been disputed. Some researchers have argued that formulas are a dead-end or side-street on the road to productive language use, whether in first- or second-language learning. Brown & Hanlon (1970) in first-language acquisition and Krashen & Scarcella (1978) in second-language acquisition observed that language learned as chunks tends to become overlearned and automatized (or fossil-

ized) and then to resist segmentation. Bates *et al.* (1988) argued that the formulaic style is not really a different route into language, but that rote output is, instead, an isolated developmental pathway which leads nowhere.

Pine (1990) has commented on the increasingly general assumption that 'expressiveness', characterized by the use of formulas, is a less advanced strategy than 'referentiality', whereas Nelson's (1973) original formulation represented only a qualitative distinction in style, not a quantitative difference in rate of development. Pine & Lieven (1990) question this view of formulas as a less advanced strategy and claim that the use of formulas by expressive children is evidence of a different rather than slower strategy than that used by referential children. They argue that a longitudinal rather than cross-sectional methodology can identify 'expressive-phrasal' children who are fast language learners.

Peters (1983, 1990) also rebuts the suggestion that formulas are not facilitative, noting that it is possible to observe formulaic breakdown feeding into the formulation of a productive system in first-language acquisition. She argues that the 'phrasal' approach of formulas CAN provide a viable route into adult morphosyntax. For example, she states:

The process of segmenting units (i.e. formulas) yields not only the sub-units but also information about the underlying structural pattern of the original unit.... Children may extract and make use of such structural patterns. (Peters, 1983: 44)

Clark (1974) observed this in data from her son Adam, and suggested that formulas acted in some cases as primary data or raw material, allowing the child to carry out a gradual analysis of their internal structure, so that they eventually became productive. Clark (1982) later characterized this as 'talking to learn'. This accords with Elbers's (1990) suggestion that the child's output may act as 'self-produced input' with newly acquired forms starting out as relatively unanalysed routines, which are later treated as input for analysis by the child. Snow (1986) also argued that there are many routes to acquisition, and that one alternative is the acquisition of formulas which the child later segments and analyses.

There appear to be two definitions of formulas co-existing in the literature on first- and second-language acquisition. On the one hand they are regarded as frozen phrases from which it is difficult to extract lexical or grammatical information; and on the other hand, as utterances which the child first uses whole and then breaks down into their components, perhaps going through an intermediate stage when only part of the unit has been analysed. The first is the 'dead-end' type which may be focused on by those who consider that formula use is evidence of a slower strategy of acquisition. The other type is the formula which the child begins to segment and analyse before it becomes fossilized, which facilitates linguistic analysis by the child. In Cruttenden's

(1981) terms, this represents the progression from 'item-learning' to 'system-learning'. He argues that language acquisition may occur initially on an 'item by item' basis, involving the learning of a form which is uniquely bonded with another form or with a unique referent. The segmentation and analysis of such an item results in 'system-learning', when the 'item' is broken down and its relationships analysed, so that it is 'built into' the child's developing system.

IDENTIFICATION OF FORMULAS

The evaluation of the role of formulas in acquisition requires that they be identified reliably and sensitively. It is not enough only to recognize the prototypical formula; the criteria must allow the consideration of more dubious instances also. Hakuta (1974) noted that the complexity of this problem means that ultimately many formulas escape identification. Nelson (1973) also commented on the intractability of the problem, opting to retain the adult definition of a word as her unit of analysis.

Ingram (1981, 1989) approached the issue from the other direction, in attempting to define productivity. He established three categories to which words were assigned: 'not lexically free', 'lexically free' and 'grammatically free'. A word was 'not lexically free' if it did not occur alone or in at least two sentences with different words. Dale (1985) (cited in Ingram 1989) adapted Ingram's (1981) analysis to allow for sentences which are 'nonproductive', in which no word is lexically free, and 'partially productive', in which at least one word – but possibly all words – are lexically free, or at least one word is grammatically free. These criteria therefore allow for different degrees of productivity. However, because they focus on distributional factors, other possible markers of non-productivity are not explicitly taken into account.

Plunkett (1990) has suggested that formulas may be recognized in early child language by their level of articulatory fluency. Following Lindblom (1985), he argues that articulatory fluency and articulatory precision are inversely related, and that the young child who has extracted a longer expression as a formula from the speech signal will produce it fluently at the expense of its articulatory precision. He noted considerable success in discriminating what he argued were formulas on the basis of their fluency and lack of articulatory precision. However, he admits that there may be a practice effect at work also, so that high-frequency productive utterances may come to be produced with these features of high fluency and low articulatory precision, and conversely, practice may result in formulaic expressions later being produced with greater articulatory precision than was initially the case. This criterion may also favour the fossilized formulas already described, rather than the formulas which are the raw material for analysis. Plunkett observes that his findings do not warrant the conclusion that articulatory/fluency criteria are necessarily a superior method for

identifying linguistic units than distributional/frequency criteria. A major difficulty with the articulatory/fluency criteria arises from the fact that they operate an on/off method of formula identification, and do not allow the stages of segmentation and analysis to be traced.

Brown (1973), Wong Fillmore (1976) and Peters (1983) proposed various criteria for the recognition of formulas in child language. Peters (1983: 8) noted that it may not be possible to make a decision for every utterance, but claimed that her set of criteria give some clues. Peters' criteria have built-in contradictions which make satisfying all of them impossible; for example, one criterion asks if the utterance is 'an idiosyncratic chunk', whereas another asks if the expression is a 'community-wide formula'. Similarly, there may be some opposition between the criterion concerning the use of the expression in inappropriate contexts, and that concerning its situational dependence. Ways of dealing with such oppositions are not discussed. Wong Fillmore made it clear that only the most clear-cut cases satisfy all her criteria, and that most others fall into a greyer zone. She noted that 'not all of the utterances classified as formulaic meet all or even most of the criteria' (1976: 312), and argued that the reader must trust the researcher's judgement. This not very satisfactory state of affairs requires further examination.

Table 1 presents a set of conditions for formula identification based on those already mentioned (with the greatest debt owed to Peters) but set in a preference rule system, which involves the interaction of a cluster of interrelated conditions as described by Jackendoff (1983). Such a system was developed to cope with 'the gradation of judgements and with the existence of exceptions to many apparently defining conditions' (Jackendoff 1983: 139). A preference rule system allows multiple converging sources of evidence for a judgement. It distinguishes between conditions which are necessary, conditions which are graded – the more something is true, the more secure is the judgement – and typicality conditions which apply typically but are subject to exceptions. Typicality conditions may reinforce or conflict with each other. Allowing exceptions does not result in a free-for-all; instead, the preference rule system imposes a control in its measure of the stability of a judgement. Jackendoff (1983: 135) points out that, although some of the rules in a preference rule system are necessary conditions, there is no subset of rules that is both necessary and sufficient, since the necessary conditions alone are too unselective. In this case, an utterance which satisfies all or most conditions receives the most stable analysis as a formula, while an utterance which fails more conditions may be judged a dubious instance, and this is made explicit. This is more sensitive than saying that 'enough' or X of the criteria must be satisfied, since the preference rule system allows a balancing of strength of confirmatory rule applications against the strength of disconfirmatory rule applications. Such an application of preference rules to the identification of formulas clarifies an area in which decision-making can

TABLE 1. *Conditions for formula identification*

Condition 1	The utterance is at least two morphemes long. (Necessary, graded)
Condition 2	The utterance coheres phonologically. (Necessary)
Condition 3	The individual elements of an utterance are not used concurrently in the same form separately or in other environments. (Typical, graded)
Condition 4	The utterance is grammatically advanced compared to the rest of the child's language (i.e. the grammatical pattern is not represented with different words). (Typical, graded)
Condition 5	The utterance is a community-wide formula, or one which occurs frequently in the parents' speech. (Typical, graded)
Condition 6	The utterance is an idiosyncratic chunk. (Typical, graded)
Condition 7	The utterance is used repeatedly in the same form. (Typical, graded)
Condition 8	The utterance is situationally dependent. (Typical, graded)
Condition 9	The utterance may be used inappropriately, either syntactically or semantically. (Typical, graded)

appear very cloudy. Some of the issues raised by this set of conditions will be discussed under the following headings: (i) length of units, (ii) frequency of occurrence, and (iii) appropriateness of use.

Length of units

Several researchers who noted the existence of formulas in second-language learners were dealing with children aged four, five and six years (e.g. Hatch, 1972; Hakuta, 1974 and Wong Fillmore, 1976). The formulas of first-language learners might be expected to be shorter, not only because of memory limitations, but also because of their less complex communicative needs. Multi-word units can be distinguished from cases where the child learns a base + inflection as an 'amalgam' (MacWhinney, 1978). However, even this minimal two-morpheme amalgam was recognized and included in Peters' (1983: 71) definition of a formula (or 'long unit') as a unit 'containing two or more morphemes of the adult language', though her discussion tended to focus on longer units.

Wong Fillmore's discussion of formulas concentrates almost exclusively on multi-word units, yet she also includes minimal units. The prototypical formula is the multi-word unit, but restricting the definition to such units

would underestimate the role of formulas in early acquisition. For this reason Condition 1 accepts two morphemes as the minimum length for an utterance to be classified as formulaic. This is a necessary condition, but also a graded condition: the longer the utterance compared to the rest of the child's language, the more likely it is to be a formula.

Frequency of occurrence

Frequent and unchanging use is generally cited as a characteristic of the prototypical formula, and it was included in Peters' and Wong Fillmore's sets of criteria. It is possible that it is the formula which occurs very frequently which is most likely to become, or have become, fossilized. Brown & Hanlon (1970) noted that frequent use of an unanalysed fragment protects it from reanalysis. On the other hand, the formula which may be most facilitative may be segmented and reanalysed early in its use. If we accept that frequency is only a typicality condition, rather than a necessary one, and may in some cases be a contra-indication of the formula which is used for data analysis by the child, then we must not rule out the possibility that an utterance which does not occur repeatedly is a formula. This would apply even to the ideal data set which contained every utterance by the child, but *a fortiori* to the more realistic studies of samples of a child's language, which might not contain more than one or two instances of a formula's use.

An example of a candidate formula which was not used frequently occurred in the data from Eibhlís, a monolingual Irish speaker aged 1;11 at the time of this example:

- (1) raghaimid síos dtí mama
 (go-Fut-we down to mama)
 V S A A
 'we will go down to mama'.

Raghaimid is the future form of an irregular verb with a synthetic first-person plural ending, and *dtí* is part of the compound preposition *go dtí* 'to'; Eibhlís did not use either *raghaimid* or (*go*) *dtí* elsewhere at this time. The utterance was produced very fluently but inappropriately, as the child's mother was sitting next to her. This was Eibhlís's only spontaneous use of the future tense of this verb in the data, and her only use of the synthetic 1st-person plural. Her mother reported that she did not use this or similar utterances outside of data collection in this period.

The complexity of this utterance compared to her mainly two-word combinations at this time, and the fluency of its production, are signs that it may have been a formula, but it was articulated precisely as well as fluently and would not have been selected by Plunkett's criterion. Because it was not used repeatedly it does not satisfy Wong Fillmore's (1976) and Peters' (1983)

requirements regarding frequent use. Bloom (1970), in discussing the identification of productive utterances, emphasized frequent and varying use of the UNDERLYING STRUCTURE rather than frequent occurrence of the utterance itself. She defined a structure as productive if it occurred five times or more 'with different formatives in different situations' (Bloom 1970: 17). Eibhlís's *raghaimid síos dtí mama* occurred at a time when she did not produce any other VSAA constructions and when her other utterances were far less complex at the two-word stage, pointing to this being a non-productive utterance for the child.

Thus, while the prototypical formula may occur frequently, this must not lead us to assume that it is ONLY utterances which occur repeatedly which may be formulas. The conditions must allow for these different types of formula, and condition 7 notes that repeated use in the same form is a typicality condition, rather than a necessary one. It is very important to bear in mind that the formula which is most significant for the child's acquisition may not occur frequently before it is broken into a formulaic frame or its components used freely.

Inappropriateness of use

Brown (1973) considered that syntactic inappropriateness was an indicator that a phrase had not been fully analysed, e.g. *it's fell*. Clark (1974) and Peters (1983) included inappropriate use in some contexts as clues to the formulaic status of an utterance. e.g. *I carry you* meaning 'you carry me'. Clark suggested that this indicated that the utterance had been given only a global interpretation with reference to the situation. Wong Fillmore suggested that 'functional overextension' of a formula supports the theory that it is not fully analysed e.g. her subject Jesus's use of *so what, you got 'em* in a variety of play situations.

Wong Fillmore and Peters also included situational dependence in their criteria for formulas, stating that an utterance which always occurs in a particular well-defined context is a strong candidate for recognition as a formula. However, an utterance which is closely tied to a particular situation generally tends to be appropriate in that situation, functionally at least. Allowing for functional overextension raises the possibility of conflict between the criteria of inappropriate use and situational dependence in some cases. It is most likely that in the later stages of development of a formula some lexical or semantic analysis does take place. A functional overextension could therefore represent a stage of hypothesis testing about the formula's function and the beginning of its analysis. To allow for this development, inappropriateness of use and situational dependence are defined as typicality conditions rather than as necessary ones, since they can conflict. Both are retained in order to accommodate the range of variation between formula types, and the development within a formula over time.

The discussion of these issues indicates that, when trying to identify formulas, an utterance must satisfy the necessary conditions, but only the most prototypical formula will satisfy most of the other conditions. This is necessary in order to make explicit the range of utterance types which may be formulas, and it allows for the fact that a formula may not remain a static entity, but may instead be in the process of analysis. The next section presents two illustrations of the application of these conditions for formula identification to data from Eibhlís, an Irish monolingual child.

The data discussed here are from a longitudinal study of the acquisition of Irish by three children. The children were taped in their homes every 18 days while interacting with a family member, usually their mothers. The data collection aimed to sample routine activities such as dressing, undressing or eating as well as play with toys and book reading.

APPLICATIONS OF THE CONDITIONS FOR FORMULA IDENTIFICATION

From 1;9 Eibhlís began to use *n'fheadar* /n'adər/ frequently. This is the contracted form of the Neg + V + synthetic first-person marker *ní fheadar* /ni: adər/ (not know-I) 'I don't know'. Eibhlís always pronounced this with a contracted negative, as adult speakers do most of the time. Her use of this form satisfies the conditions as follows.

- (1) The utterance is composed of more than two morphemes.
- (2) It cohered phonologically.
- (3) While the uncontracted negative *ní* was used in other environments, there were no other uses of the contracted form, or of *fheadar*.
- (4) It was grammatically advanced compared to the rest of the child's language at 1;9. When Eibhlís began using *n'fheadar* she did not use synthetic person markers on any other verb. Her infrequent use of *ní* occurred with verb-less constructions, or alone so that the structure Neg V S was confined to this representation.
- (5) The utterance is a community-wide formula, and occurs frequently in the parents' speech. However, adults also use this (defective) verb in questions such as *an bhfeadaraís?* 'do you know', and with the third-person pronoun *ní fheadair sé* 'he doesn't know', whereas none of these uses occurred in Eibhlís's data.
- (6) The utterance was not an idiosyncratic chunk.
- (7) It was used frequently in the same form.
- (8) It was situationally dependent, insofar as it was used only as a reply to *wh*-questions.
- (9) There was no syntactic overgeneralization of the form, but it was used inappropriately at times, when it appeared to have been interpreted as a general negative response indicating rejection, e.g.

- (2) D. conas táin tú Eibhlís? ('how are you Eibhlís?')
 E. n'fheadar (*rejecting his attempt to initiate their routine based on this question.*)

Thus it appears that Eibhlís's use of *n'fheadar* satisfies the necessary conditions, and most of the typicality conditions for formula identification. This formula was used very frequently throughout the rest of the data collection.

This raises the question of whether any analysis occurred on this formula while it continued to be used. The persistence of *n'fheadar* does not mean that Eibhlís did not also develop a productive negative, or analyse synthetic person markers. By 2;1, she was using Neg+V phrases such as *ní thug* (Neg give-Past), and she used other pronouns and synthetic person markers, though the verb *feadair* only ever occurred in this environment in her data (this verb can only occur in negative or interrogative constructions; a different construction is used in Irish to indicate when something is known: *tá a fhios agam* 'be knowledge at-me').

A related question concerns whether *n'fheadar* could be considered a formula after at least one of its elements had begun to be used separately and in other environments. Garman's (1987) data on his son Christopher's use of *it's coming* and *here it comes* suggests that it is possible that a child could use the same morpheme in different environments and nevertheless not have analysed it as the same. Brown (1973) gave examples of this in his data, where even when there was clear evidence that *be* had been correctly organized by Adam, *it's* persisted as an amalgam. Brown also experienced this in his own learning of Japanese, when he mis-segmented *kore wa* as one word, and did not correct it even when he learned that *wa* is a particle which he also used in the environments *sore wa* and *are wa*. It was only some time later, when he learned *kore o* that he fully analysed *wa*.

Thus, the concurrent production of the same morpheme in different environments does not necessarily mean that it has been completely analysed and is productive. It may not have been matched as the same element in two formulas, or it may be the unanalysed part of a formulaic frame. However, there must be a limit on the degree of freedom to occur in the same form in other environments. For example, it would be difficult to argue that *n'fheadar* was completely unanalysed if both the contracted negative and *fheadar* occurred separately and freely in other environments.

On the other hand, it is also possible, as Peters (1983: 79) suggests, that for speed of processing, even fully analysed units may be used as if they were unitary, so that the learner may for a time retain in the lexicon both the large chunk and its elements. Bolinger & Sears (1981: 53) note that 'these whole chunks that we learn persist as coded units even after analysis into words has partially split them up.' However, in the estimation of MLU such utterances

would not be counted as single units in either adult or child language, once it was clear that all of their morphemes had been segmented and analysed. Thus Condition 3 points out that, the more elements of an utterance occur in the same form either separately or in different environments, the less likely the utterance is to be a formula. This is coded as a typicality condition to allow for the situation described by Brown & Hanlon (1970: 51) in which a fossilized formula such as *what's that?* which had resisted segmentation was not reconstrued in terms of the new analysis of other productive *wh*-questions. As Plunkett (1990) points out, it is possible that a morpheme which may be imprecisely articulated as part of a longer unit may not be recognized as the same morpheme which the child uses productively in other environments, so this condition is not a necessary one.

Application to a formulaic frame

An example of a phrase that moved from being formulaic to productive through a stage as a 'formulaic frame' can be seen in Eibhlís's *dom é* (to-me it), which she used between 1;7 and 1;9 as a demand, the full form of which is *tabhair dom é* (give to-me it) 'give me it'. Applying the preference conditions to this utterance at 1;7 we find that:

- (1) The utterance was at least two morphemes long.
- (2) It cohered phonologically.
- (3) *dom* (to-me) was not used in any other environment, but the pronoun *é* had occurred between 1;4 and 1;6 in *cad é?* (what-is it).
- (4) It was a construction unrelated to any productive pattern in the child's language at that time. Eibhlís used no other Adverb-Object utterances during this period, nor any other prepositional pronouns such as *dom*.
- (5) It was not a community-wide formula, but was used reasonably frequently in her home, either by her mother or her brother.
- (6) It was an idiosyncratic chunk.
- (7) It was used repeatedly in the same form.
- (8) It was situationally dependent, used when she wanted something handed to her.
- (9) It was not used inappropriately.

Table 2 summarizes the development of this phrase in Wong Fillmore's format, where S[F] indicates a formula, and S[Fx]¹ and S[Fx]² are formulaic frames.

Eibhlís began at 1;7 with the formula, *dom é*. When her mother tried to elicit an imitation of the full phrase at this time, by stressing the imperative, Eibhlís imitated only the verb *tabhair* 'give' and did not use it spontaneously. A directed imitation without the stress on the verb in the same session

TABLE 2. *Eibhlís: Development of tabhair dom é 'give it to me'*^a

Age 1;7		
S[F] dom é	dom é (× 12)	(to-me it)
Imitation	dom é	(M. Tabhair dom é (give to-me it) 'Give it to me' (<i>attempt to elicit tabhair</i>) (give) (M. TABHAIR dom é 'GIVE me it') (M. Abair tabhair dom é 'Say give me it')
Directed imitation:	tabhair	
Directed imitation:	dom é	
Age 1;9		
S[F] dom é	dom é (× 2)	
S[Fx] ¹ dom X	dom an potty	(to-me the potty)
X = NP	dom an nappy	(to-me the nappy)
	dom an báidí	(to-me the boat)
	dom an leabhar	(to-me the book)
Imitation:	tabhair bat di	(M. Tabhair bat di 'Give her a bath')
Age 1;10		
S[Fx] ² tabhair dom X		
X = NP	tabhair dom é	(give to-me it)
	tabhair dom biro	(give to-me a biro)
Directed imitation:	tabhair	(M. Abair tabhair dom 'Say give me')
Imitation:	tabhair dom é (× 2)	(M. Tabhair dom an buidéilín abair. 'Give me the bottle say')
Age 2;0		
tabhair is freed:		
	tabhair	
	*tabhair capaillín	(give a-horsie)
	*tabhair an ceann eile	(give the one other)
	*tabhair capaillín dtús	(give a-horsie first)
	*tabhair mise ceann eile	(give I one other)
	tabhair domsa leabhar eile	(give to-me book another)
	tabhair dom é (× 2)	(give to-me it)
	tabhair di an baba	(give to-her the doll)
Expanded imitation	*tabhair póigín	(M. Tabhair gráin di)
	'give a kiss'	'Give her a love'
Expanded imitation	tabhair hubí di	(M. Tabhair póigín di)
	'give her a lift-up'	'Give her a little kiss'
S[F]remains	dom é (× 5)	(to-me it)
S[Fx] ¹	dom é sin (× 5)	(to-me that)
	mamaí dom é	(mamaí to-me it)

* Indicates an utterance unacceptable to an adult.

^a S[F] = formula; S[Fx]¹, [S[Fx]]² = formulaic frames.

yielded only *dom é*. Between 1;9 and 1;10 she began to specify the referent, e.g.

(3) dom an nappy (to-me the nappy);

(4) dom an báidí (to-me the boat)

thereby treating *dom X* as a frame on which to hang NPs. During this period other forms of the prepositional pronoun paradigm in which *dom* is the first person were not used spontaneously, nor was *dom* used in any other

construction. She sometimes used what may have been a place holder before the prepositional pronoun, inserting a schwa instead of the still missing imperative *tabhair*. A spontaneous imitation at 1;9 included the imperative verb and the third-person pronoun *di* instead of the first person *dom*. By 1;10 Eibhlís was using the entire phrase *tabhair dom é*.

At 2;0, *tabhair* was freed from the rest of the phrase, as evidenced by several examples in the table, among them the following:

- (5) M. dtabharfaimíd Tina linn? ('Will we bring Tina with us?')
E. **tabhair* (bring 'yes.')

However, from this point Eibhlís was faced with creating her demands from scratch whenever she used the imperative, and this led to errors such as the following ill-formed utterances which lack an indirect object:

- (6) **tabhair capaillín* (give a-horsie)
**tabhair an ceann eile* (give the one other)
**tabhair capaillín dtús* (give a-horsie first)
**tabhair mise ceann eile* (give I-emphatic one other)

Not all *tabhair* requests were ill-formed, however; when she reverted to her formulaic frame she produced *tabhair dom é* correctly several times, and once even *tabhair domsa leabhar eile* (give to-me-emphatic book other) 'give ME another book' (although the emphatic was not appropriate in the context). Her frame *dom X* continued to be used at this time as often as her newly formed *tabhair X* utterances, but it seemed that once *tabhair* had been isolated as the verb it lost one of its necessary arguments. The error **tabhair mise ceann eile* (give I-emphatic one other) at age 2;0 indicates that she did still use 'give' as a two-argument verb, but without using her formula she was not always able to express it correctly. It also indicates that her incomplete *tabhair X* utterances were not the result of a length constraint, since this ill-formed utterance is longer than her well-formed ones. Neither can it simply be the result of using long and unfamiliar NPs as object, since the very frequently used *capaillín* 'little-horse' also occurred in an incomplete *tabhair X* utterance.

Tracing the progression from such a formula through its different levels of segmentation and analysis to productivity allows a more sensitive evaluation of the contribution of formulas to the acquisition process than an analysis which assumes that an utterance is either fully productive or not. There are bound to be some cases where distinguishing between a formulaic frame and a productive utterance is difficult, but the attempt to do so is more defensible than assuming that all such utterances are fully productive.

CONCLUSION

There are difficulties in identifying formulas in young children's data, but these are not insuperable. What is very necessary is a statement of the criteria used in establishing that a particular utterance constitutes a formula. The use of a preference system makes clear the distinction between necessary conditions and typicality conditions in formula identification. It highlights the fact that there are sub-types of formula, which can be manifested differently, but which may have a cluster of characteristics in common. It also attempts to deal with the oppositions implicit in the criteria proposed by other researchers. If the contribution of such formulas to the acquisition process is not to be underestimated, then there is a need for greater clarification of the procedures used in identifying utterances as formulas, and a closer examination of the utterance's history to trace its development towards productivity.

REFERENCES

- Bates, E., Bretherton, I. & Snyder, L. (1988). *From first words to grammar: individual differences and dissociable mechanisms*. Cambridge: C.U.P.
- Bloom, L. (1970). *Language development: form and function in emerging grammars*. Cambridge, MA: M.I.T. Press.
- Bloom, L., Lightbown, P. & Hood, L. (1975). Structure and variation in child language. *Monographs of the Society for Research in Child Development* 40, (Serial No. 160).
- Bolinger, D. (1976). Meaning and memory. *Forum Linguisticum* 1, 1-14.
- Bolinger, D. & Sears, D. (1981). *Aspects of language*. (3rd edn) New York: Harcourt Brace Jovanovich.
- Bretherton, I., McNew, S., Snyder, L. & Bates, E. (1983). Individual differences at 20 months: analytic and holistic strategies in language acquisition. *Journal of Child Language* 10, 293-320.
- Brown, R. (1973). *A first language: the early stages*. London: Allen & Unwin.
- Brown, R. & Hanlon, C. (1970). Derivational complexity and order of acquisition in child speech. In J. R. Hayes (ed.), *Cognition and the development of language*. New York: Wiley.
- Clark, R. (1974). Performing without competence. *Journal of Child Language* 1, 1-10.
- (1982). Theory and method in child language research: are we assuming too much? In S. Kuczaj (ed.), *Language development*, Vol. I: *Syntax and semantics*. New Jersey: Erlbaum.
- Cruttenden, A. (1981). Item learning and system learning. *Journal of Psycholinguistic Research* 10, 79-88.
- Dale, P. (1985). Productivity analysis of early child language (PANEL): user's guide. IBM versions. Unpublished paper.
- Elbers, L. (1990). Language acquisition: blends, overgeneralizations and self-produced input. In P. Coopmans, B. Schouten & W. Lonneveld (eds), *OTS Yearbook 1990*. Dordrecht: ICG.
- Garman, M. (1987). The role of linguistics in speech therapy: from assessment to remediation. In P. Grunwell & A. Jones (eds), *The functional evaluation of language disorder*. London: Croom Helm.
- Hakuta, K. (1974). Prefabricated patterns and the emergence of structure in second language learning. *Language Learning* 24, 287-97.
- Hampson, J. & Nelson, K. (1990). Early relations between mother talk and language development: masked and unmasked. *Papers and Reports on Child Language Development* 29, 78-85.

- Hatch, E. (1972). Some studies in language learning. *UCLA Workpapers in Teaching English as a Second Language* 6, 29-36.
- Ingram, D. (1981). Early patterns of grammatical development. In R. Stark (ed.), *Language behavior in infancy and early childhood*. New York: Elsevier North-Holland.
- (1989). *First language acquisition: method, description and explanation*. Cambridge: C.U.P.
- Jackendoff, R. (1983). *Semantics and cognition*. Cambridge, MA: M.I.T. Press.
- Jespersen, D. (1924). *The philosophy of grammar*. (Reprinted 1965.) New York: Norton.
- Krashen, S. & Scarcella, S. (1978). On routines and patterns in language acquisition and performance. *Language Learning* 28, 283-300.
- Lindblom, B. (1985). Phonetic universals in vowel systems. In J. J. Ohala (ed.), *Experimental phonology*. San Francisco: Academic Press.
- MacWhinney, B. (1978). The acquisition of morphophonology. *Monographs of the Society for Research in Child Development* 43, Serial no. 174.
- Nelson, K. (1973). Structure and strategy in learning to talk. *Monographs of the Society for Research in Child Development* 38, Serial no. 149.
- (1981). Individual differences in language development: implications for acquisition and development. *Developmental Psychology* 17, 170-87.
- Peters, A. (1983). *The units of language acquisition*. Cambridge: C.U.P.
- (1990). The morphosyntactic development of a phrasal child. Paper presented at the Fifth International Congress for the Study of Child Language, Budapest.
- Pine, J. (1990). Non-referential children: slow or different? Paper presented at the Fifth International Congress for the Study of Child Language, Budapest.
- Pine, J. & Lieven, E. (1990). Referential style at 13 months: why age-defined cross-sectional measures are inappropriate for the study of strategy differences in early language development. *Journal of Child Language* 17, 625-31.
- Plunkett, K. (1990). The segmentation problem in early language acquisition. *Center for Research in Language Newsletter* 5, No. 1.
- Snow, C. (1986). Conversations with children. In P. Fletcher & M. Garman (eds), *Language acquisition* (2nd edn) Cambridge: C.U.P.
- Wong Fillmore, L. (1976). The second time round: cognitive and social strategies in second language learning. Unpublished doctoral dissertation, Stanford University.