
**The Beginning of Morphological Learning: Evidence from Verb Morpheme Processing in Preverbal Infants**

**Alexandra Marquis & Rushen Shi**

**Abstract**

The goals of our chapter are to study infants’ acquisition of verb morphology and develop a model of early morphological learning. We approach this acquisition by studying French-learning infants’ initial segmentation of verb forms and their interpretation of verb morphological alternations. Our empirical findings demonstrate that infants begin to parse verbs into decomposed stems and suffixes by 11 months of age, and they have rudimentary knowledge of regular verb paradigms by 14 months of age. We show that this learning is based on distributional analyses of the input without the need for semantics.

**Keywords**: Morphological learning; Verb morphology; Language acquisition; Infant speech processing; Function words/morpheme; Bootstrapping; French; Cognitive development

**1. Introduction**

Verbs in many languages are believed to be harder to acquire than nouns. This is in part because in these languages, verbs bear a high load of information via morphological affixes entailing tense, aspect, number, etc. Affixes are bound morphemes that never surface alone without a root or stem. Young children who have not yet acquired their language must thus cope with a multitude of forms for the same verb (e.g., *call, calls, called, calling*). Moreover, morphological affixations often modify the syllable structures of verbs (e.g., the sound */l/* is
the coda in the monosyllabic root *call* */kɔ:l/, but becomes the onset of the second syllable in *calling* */kɔ:liŋ/), rendering the recognition of stems and affixes more difficult. Two competing views may be considered regarding verb representations. Decomposed roots and suffixes (e.g., *call*, *-s, -ed, -ing*) may be stored as separate pieces. Stored pieces can be later linked together at the level of morphological paradigms. Or, in a non-decomposition view, variable forms of a verb may be encoded in the lexicon as single units (e.g., *call, calls, called, calling*) without the segmentation of the suffixes. The difference of these two views has important implications for the understanding of initial morphological development in infants. The non-decomposition view is consistent with the assumption that infants may not be capable of parsing affixes and stems, nor does the grammar require them to do the parsing. On the contrary, the decomposition view predicts that infants can segment stems and affixes. Furthermore, on this view, a verb stem in the unaffixed bare form and in affixed forms may be treated as variants of the same unit despite considerable differences due to phonological operations.

Obviously, the full knowledge of morphological paradigms involves the acquisition of meaning for stems and affixes. A crucial question concerns the initial state of morphological learning: do infants rely on semantics to break into the learning of morphological paradigms? That is, do infants first have to learn word meaning in order to understand the relatedness of bare/unaffixed and affixed forms of a verb? Or do they instead begin learning morphological relatedness of word forms by distributional analyses without relying on semantics? We suggest that the latter is more plausible, especially for verb learning, given the evidence that infants have more difficulty learning verb meaning at the early stage of acquisition (e.g., Gentner, 1982). Although infants’ very first encoded pieces (e.g., stems and suffixes, or non-segmented affixed verbs, depending on the theory) may be achieved through successful
analyses of the input without semantics in both decomposition and non-decomposition theories, different predictions can be made about subsequent morphological learning.

In this paper, we argue for the decomposition view for infants’ morphological development. In particular, we suggest that infants at the initial learning stage parse verb stems and affixes without relying on semantics, but on the basis of high token frequency of affixes and high type frequency of stems (i.e., regular morphological operations). The encoding of affixes is likely, given the cross-linguistic evidence that preverbal infants from six months of age begin to recognize highly frequent function words such as determiners (e.g., Hallé, Durand, & de Boysson-Bardies, 2008; Höhle & Weissenborn, 2003; Shafer, Shucard, Shucard, & Gerken, 1998; Shi, Cutler, Werker, & Cruickshank, 2006; Shi & Lepage, 2008; Shi, Werker, & Cutler, 2006), long before they understand the meaning of these words. In addition, we suggest that infants can perceive a stem in different morphological contexts as variants of the same abstract form. Moreover, this knowledge can bootstrap infants into the learning of verb meaning, in the sense that the learning system expects the bare stem and segmented stems from affixed contexts to have the same core meaning, as it would when acquiring the meaning of a non-affixed word from different utterance locations. Thus, in our view, infants’ earliest stem-affix segmentation already represents rudimentary morphological knowledge, which may facilitate the subsequent learning of verb meaning, which in turn may lead to the generalization and abstraction of full morphological paradigms. This position contrasts widely with theories that require infants to first learn verb meaning in order to learn the morphological relations shared among unaffixed and various affixed forms of a verb. The non-decomposition account, for example, would need semantics for infants to understand the link between the wholly encoded unit calling and the unaffixed form call.

There is evidence from perceptual studies that by six months of age, infants already begin to recognize highly frequent word forms without knowing their meaning (e.g., Bortfeld,
Morgan, Golinkoff, & Rathbun, 2005; Shi, Marquis, & Gauthier, 2006). Previous research showed infants’ capacities to perform statistical analysis of speech (e.g., Johnson & Jusczyk, 2001; Saffran, Aslin, & Newport, 1996). In these studies, infants were capable of segmenting word-like units based on the statistical distribution of syllables of a novel language. What remains to be tested is whether infants can perform such distributional analysis on subsyllabic units that correspond to bound morphemes such as suffixes. In our research, we seek to answer this question, and further, to determine whether the outcome of suffix parsing represents rudimentary understanding of morphology.

The question of morphological parsing is tied to the question of embedded words. In the case of morphologically varying forms such as call-calling /kɔːl-/’kɔːlɪŋ/, the first part of calling (i.e., call-) and the bare form call should be interpreted as the same abstract unit. Non-morphological words involving embedded forms cannot be interpreted in the same way. For example, sir /sɜːr/ and circle /sɜːrkəl/ share no internal relation. Natural languages must require mechanisms that balance between these different linguistic functions: relatedness for morphosyntactic needs and unrelatedness for increasing vocabulary contrasts (e.g., adding phonemes or a syllable to create new words, sir-circle /sɜːrkəl/). We suggest that frequency is the mechanism that makes this balance work: high token frequency of the suffix co-occurring with high type frequency of the stem leads to stem-suffix segmentation. For example, the –ing suffix in English occurs highly frequently with a large number of different stems (e.g., calling, walking, eating, drinking, dozing, leaving, etc.), allowing –ing and stems to be segmented. Furthermore, infants operate with a bias that allows them to expect the segmented stems (e.g., call- from calling, called) and the bare stem (e.g., call) of a verb to be the same unit and to later receive the same meaning. This bias is
governed by frequency conditions and is inhibited in the case of *sir-circle* since the frequency requirement is not met.

Existing research on infants’ interpretation of embedded words provide some indirect supporting evidence. Jusczyk and colleagues (Jusczyk, Houston, & Newsome, 1999) found that English-acquiring infants did not accept *ham* as a variant of *hamlet*, suggesting that they encoded the disyllabic noun as a whole, and perceived the words *ham* and *hamlet* as unrelated. Although infants at 8 months of age sometimes mis-segment embedded forms such as *tar* in *guitar* due to prosodic factors, by 10-11 months they can overcome the weak prosody of *gui-*, perceive the distributional integrity of the two syllables, and treat the disyllabic word *guitar* and the monosyllabic word *tar* as different forms. Yet, perception of embedded forms in morphological cases presents a different story. A study on verb segmentation (Mintz, 2013) showed that English-learning infants seem to understand the link between a novel bare stem and the stem plus the present progressive -*ing* (e.g., *lérjov* and *léřjoving*). However, because infants in Mintz’ study (2013) were familiarized with suffixed forms and tested with the bare stem versus a novel word, results can also be interpreted as indicating that infants only mapped the form of the test word with the beginning of the familiarized words without necessarily processing the forms in a morphological fashion. A study in which infants are directly tested on their stem representation is thus needed. Such a study will need to present a suffixed form and test infants on both possible parses (i.e., stem versus partial stem). In this paper we present our recent empirical data that demonstrate that infants not only associate suffixed and unsuffixed verb forms, but they also encode detailed representations for word-internal morphemes in a way that reflects rudimentary knowledge of morphological alternations.

2. The beginning of verb parsing: Experiment 1
We used French, a language with substantial verb morphology, as the test case for examining early morphological learning. Before investigating infants’ processing of verb form alternations, we first needed to determine the age at which they begin recognizing unaffixed verb forms in sentential contexts. One verb segmentation study with infants (Nazzi, Dilley, Jusczyk, Shattuck-Hunagel, & Jusczyk, 2005) has been previously published, showing that English-learning infants can segment bisyllabic verbs (e.g., discount, permit) starting from 13.5 months of age. We hypothesized that infants may segment monosyllabic verbs at a younger age. We thus tested 8- and 11-month-old French-learning infants (16 for each age group) using a visual preferential procedure (Cooper & Aslin, 1990), a procedure that has been used in previous infant segmentation studies (e.g., Curtin, Mintz, & Christiansen, 2005; Shi & Lepage, 2008). In this experiment (Marquis & Shi, 2008), infants were familiarized with trials presenting repeatedly a CVC unaffixed verb form, either /bif/ or /tar/, until they reached 30 s of total looking time while listening to the target. Following the familiarization phase, infants were tested with trials of sentences containing the bare verb form /bif/ versus trials of sentences containing the bare verb form /tar/. Note that such bare forms appear in most present tense conjugations, except for 1st and 2nd person plural, and they appear in the singular imperative. We expected that if infants recognized the target verb, they should discriminate trials containing the target and those not containing the target during the test phase.

During the experiment, the infant sat on his or her parent’s lap in front of a monitor and a loudspeaker in an acoustic chamber. The parent heard masking music through headphones. An observer in the adjacent room observed the infant and pressed down a computer key whenever the infant looked towards the monitor. A computer program (Cohen, Atkinson, & Chaput, 2000) presented audio-visual stimuli, calculated the infant’s looking

---

1 Note that all infants in this paper were exposed to French for minimum of 70% of their awake time.
times online, and controlled the progression of the experiment from the familiarization to the test phases. Each trial was initiated by the infant’s look towards the monitor. Between trials, a flashing light appeared on the monitor to attract the infant’s attention.

Infants were randomly assigned to either /bif/ or /tar/ familiarization condition. The first test trial presented either the /bif/ or /tar/ sentences, counterbalanced across infants, and the two types of test trials were presented in alternation for a total of 10 test trials (see Appendix 1). The target verbs /bif/ ‘to cross out’ and /tar/ ‘to weigh’ are infrequent in spoken French (Beauchemin, Martel, & Théoret, 1992) thus probably unknown to young children. Their low frequency will allow us to assess not only infants’ ability to segment unaffixed verb roots, but also their ability to segment novel verbs. Stimuli were recorded by a native Quebec-French female speaker in infant-directed speech style.

Our prediction was that if infants were capable of recognizing the unaffixed verb roots, they should pay greater attention to the sentences containing the familiarized target than to those containing the other novel verb during the test phase. We expected a looking preference for the target sentences given that previous published infant segmentation studies using natural speech stimuli typically showed looking preferences for targets (e.g. Jusczyk & Aslin, 1995; Jusczyk et al., 1999; Polka & Sundara, 2003; Shi & Lepage, 2008). We also predicted that the 11-month-olds, having had more experience with their native language, should perform better that the 8-month-olds in segmenting unaffixed verbs from continuous speech.

Results confirmed our predictions. Average looking time per trial for each trial type during the test phase was calculated, and the data were analyzed in a 2X2 mixed analysis of variance, with Sentence Type (sentences containing the familiarized target verb versus sentences containing the non-familiarized verb) as the within-subject factor, and Age (8 versus 11 months) as the between-subject factor. We obtained a Sentence Type x Age
interaction, $F(1,30) = 6.41, p = .017$, and no main effect of Sentence Type nor Age, indicating that the two age groups differed. In follow-up $t$-tests, the group of 11-month-olds looked significantly longer during target trials than during non-target trials, paired $t(15) = 2.284, p = .037$ (see Figure 1), revealing that the 11-month-olds recognized the target verbs. In contrast, the 8-month-olds listened equivalently to both types of test trials, paired $t(15) = 1.147, p = .269$, thus showing no evidence of recognizing unaffixed verb forms in sentences. All $t$-tests reported in this article are 2-tailed.

These results (Marquis & Shi, 2008) show that by 11 months of age, infants are capable of segmenting unaffixed verb roots from continuous speech. Moreover, the fact that they were capable of segmenting these infrequent verbs suggests that our assumption is correct about semantics being unnecessary for infants’ early speech processing. The next step was to test whether infants are capable of segmenting verb stems from suffixed forms and associating suffixed verb forms with its corresponding unsuffixed stem.

3. Infants’ parsing of verb stems and suffix: Experiment 2

We established in Experiment 1 that French-learning infants are capable of segmenting unaffixed verb stems from continuous speech stream as early as 11 months of age. We then set out to examine whether infants at this age can segment affixed verbs into smaller morphemic pieces, stem and suffix, as predicted in the decomposition theory. In Experiment 2 (see Marquis & Shi, 2012), French-learning 11-month-old infants were familiarized to a nonce CCVC form, either /glyt/ (phonetically realized as [g]Yt) or /trid/ (phonetically realized as [tr]d), until they reached 30 s of total looking time. Following familiarization, infants were tested with the same procedure as in Experiment 1. One group of 16 infants were tested with Condition 1: sentences containing [gly.te] versus sentences containing [tri.de] (i.e.,
the CCVC forms conjugated with a French verb morpheme, the suffix /el/, which is the form of an infinitive, a past participle, or the second person plural present/imperative forms) (see Appendix 2). These morphemes (i.e., the infinitive, the past participle and second person plural present/imperative) are homophonous, such as the infinitive ending in *manger* /mā.ʒe/ ‘to eat’, the past participle ending in *mangé* /mā.ʒe/ ‘eaten’, and second person plural present/imperative *mangez* /mā.ʒe/ ‘eat/youPL eat’. The other group of 16 infants was tested with Condition 2: sentences containing [gly.tu] versus sentences containing [tri.du], both of which are disyllabic mono-morphemic nonsense words (see Appendix 3). This /u/ vowel exists in the French language but is not used as a verb morpheme, thus we could use this nonsense suffix as a control to directly test infants’ morphological knowledge of /e/.

Nonsense verbs were used in our experiments for two reasons: 1) to ensure that all forms were novel to infants so that we could reliably assess their generalized ability to perform morphological parsing when encountering unknown verbs; 2) the first vowel of these forms is subject to vowel alternations, [glYt]-[gly.te]-[gly.tu], [trId]-[tri.de]-[tri.du]; according to Quebec-French phonological rules, tense high vowels are laxed when they appear in closed syllables (e.g., Walker, 1984); we were therefore able to test whether infants have the ability to segment underlying stems and the suffix /e/ despite such surface changes, an ability beyond simply mapping the bare root form of the familiarization phase with part of the suffixed forms of the test phase. The same native French speaker as in Experiment 1 recorded the new stimuli.

Infants were randomly assigned to either [glYt] or [trId] familiarization condition. After familiarization, the first group of infants was tested with [gly.te] sentences versus [tri.de] sentences in alternating trials, and the second group was tested with [gly.tu] sentences versus [tri.du] sentences, for a total of 14 trials for each test group. The counterbalancing followed the same steps as in Experiment 1.
We predicted that during the test phase, if the first group succeeded in processing the vowel alternation rule in a morphological fashion, that is, recognizing the frequent /e/ verb suffix, segmenting the underlying form of the stem and matching it with the familiarization target, they should then prefer to listen to the test sentences containing the suffixed target verb. On the other hand, a failure in this experiment could support the non-decomposition view according to which bare stem forms and affixed forms are separately stored as whole units without affix segmentation.

Furthermore, if infants are truly able to segment word-internal stems and suffixes, the group of infants tested with sentences containing [gly.tu] versus those containing [tri.du] mono-morphemic forms (i.e., the targets ending with the non-morpheme /u/) should demonstrate no preference. That is, the [gly.tu] and [tri.du] forms present in the test sentences should be treated as unrelated to the familiarized [glYt] and [trId] forms. This interpretation would apply only if the first group (the /gly.tel-/tri.de/ test condition) showed a significant looking preference for the targets.

Infants’ average looking time for each trial type during the test phase was calculated, and the data were analyzed in a 2X2 mixed ANOVA with Sentence Type (sentences containing the familiarized target root versus sentences containing the non-familiarized root) as the within-subject factor and Condition (Condition 1 real verb suffix /e/ versus Condition 2 non-morphemic /u/) as the between-subject factor. This comparison yielded a significant interaction of Sentence Type x Condition, $F(1,30) = 4.836, p = .036$, while no other main effect was obtained (see Figure 2). As predicted, infants tested with the targets ending with the real suffix /e/ looked significantly longer while listening to the sentences containing the affixed form of the familiarized root in comparison to the sentences containing the novel suffixed verb, paired $t(15) = 3.113, p = .007$. This could be interpreted as evidence for the decomposition of the targets as /glyt + e/ and /trid + e/. Importantly, for the group tested with
the targets ending with the nonce suffix /u/, looking times for the sentences containing the familiarized root and those containing the non-familiarized root were not significantly different, paired \( t(15) = .945, p = .359 \), demonstrating that infants in Condition 2 did not decompose the disyllabic forms into /glyt + u/ and /trid + u/. Had infants just been parsing any initial overlapping sound sequences in this experiment without attending to the morphological status of word endings, a preference for target sentences obtained with the first group of infants should have also been found with the second group.

[Insert Figure 2 here]

Taken together, these results (Marquis & Shi, 2012) favor the decomposition view of early morphological learning. The novel verbs were parsed only in cases involving a real suffix, as /glyt/, /trid/, /e/\). The mono-morphemic forms /gly.tu/ and /tri.du/, which have a phonologically comparable structure as /gly.te/ and /tri.de/, were not parsed into smaller units. Infants in Condition 2 treated the word forms /glyt/ and /gly.tu/ (and /trid/, /tri.du/) as unrelated lexical items.

4. Rudimentary knowledge of morphological paradigms: Experiments 3 and 4

The empirical data from Experiment 2 suggest the decomposition of stems and affixes by French-learning infants at 11 months of age. One question remains concerning whether our results unambiguously indicate the knowledge of a morphological paradigm, i.e., whether infants represent /glyt/-/gly.te/ as alternations of the same abstract verb. It is possible that infants in Experiment 2 segmented the suffixed verb, but only represented the initial phonemes of the stem with phonetic details, leaving the final phoneme or phonemes unspecified. This interpretation is possible because the test comparison trials involved a completely novel suffixed verb bearing no phonological resemblance to the familiarized stem (i.e., /tri.de/ after /glyt/, or /gly.te/ after /trid/). The same results could have been obtained if
infants only recognized the frequent suffix /e/ and the initial part of the affixed verbs. It would then be uncertain whether infants can fully represent verb form alternations. We therefore decided to conduct a more direct test of the knowledge of morphological paradigms.

In the new experiments, we familiarized infants with a suffixed nonsense verb /gla.te/ and tested them with the stem form /glat/ versus a partial stem /gla/. This design forced infants to make a decision about which form during the test phase was related to the familiarized form /gla.te/. If infants had rudimentary knowledge about the verb paradigm involving the highly frequent and regular /e/ suffix, they should then segment /e/ and recognize /glat/ as the variant related to /gla.te/. The partial stem /gla/ should be regarded as having no relation with /gla.te/ although the two forms overlap at the initial portion. A control condition was designed, which included a nonsense form /gla.tu/ as the familiarization form (note again that /u/ is not a suffix in French), and also /glat/ versus /gla/ trials during the test phase. In this control condition, we predicted that infants should perceive /gla.tu/ as a disyllabic mono-morphemic word and regard the two forms presented in the test phase as equally unrelated to /gla.tu/. This interpretation of null results would hold only if the results of the experimental condition involving /gla.te/ turned out to be the pattern that we predicted. Thus, for the current experiment, we created new non-words, different from those used in Experiment 2. The vowel alternation was removed here (/a/ does not alternate). That is, the vowel of the target and test words remained constant. Because detailed knowledge of verb paradigms is more advanced than suffix parsing, we deliberately avoided vowel changes so as to better determine whether infants understood verb alternations when both forms in the test phase have exact partial phonetic overlap with the target form in the familiarization phase. The same speaker recorded the new stimuli as in the previous experiments.

In Experiment 3, 11-month-old Quebec-French-learning infants (16 per condition) were familiarized with a nonsense word, either /gla.te/ (affixed) or /gla.tu/ (mono-
morphemic), during 33 s. All infants were tested with /glat/ versus /gla/ trial types. Multiple exemplars of each word were presented. The first test trial was either /glat/ or /gla/, counterbalanced across infants. If at this age infants understand morphological alternations, then a looking preference for the stem form /glat/ should be observed only for the infants familiarized with the suffixed form /gla.te/. For the infants familiarized with the mono-morphemic form /gla.tu/ (ending with the non-suffix /u/), two outcomes were possible: 1) no preference for either test words (see the discussion of the interpretation of this control condition in the above paragraph); or 2) a preference for /gla/, suggesting a syllabic bias for parsing that was unrelated to morphological processing.

Infants’ looking times during the test phase were analyzed in a 2X2 mixed ANOVA, with Parse Type (/glat/ versus /gla/) as the within-subject factor and Familiarization (/gla.te/ versus /gla.tu/) as the between-subject factor. Results revealed a significant interaction of Parse Type x Familiarization, $F(1,30) = 4.194, p = .049$, while no other main effect was obtained. For the group of infants familiarized with the real French suffix /e/ (i.e., /gla.te/), even though their looking times appear to be longer for the /glat/ stem trial ($M = 7.438$ s, SE = 1.242 s) than for the /gla/ partial stem trial ($M = 5.225$ s, SE = .650 s), this difference did not reach significance, paired $t(15) = 1.522, p = .149$. As for the group of infants familiarized with the nonce suffix /u/, looking times for the /glat/ ($M = 5.113$ s, SE = 1.128 s) and the syllabic /gla/ trial types ($M = 6.663$ s, SE = 1.070 s) were also not significant, paired $t(15) = 1.379, p = .188$. Therefore, there was no clear evidence that the 11-month-old infants in this experiment stored affixed verb forms as decomposed stem and affixes, although they begin to show a tendency towards this direction. We thus pursued our investigation with 14-month-olds, an older group of French-learning infants.

In Experiment 4 (Shi & Marquis, 2009), 14-month-old infants (16 per condition) were familiarized and tested in exactly the same way as those in Experiment 3. It was predicted that
by 14 months of age infants should demonstrate more robust morphological knowledge. The looking responses during the test phase of the /gla.te/ (affixed) and the /gla.tu/ (non-affixed) familiarization groups were analyzed in the same 2X2 mixed ANOVA as in Experiment 3, with Parse Type (/glat/ versus /gla/) as the within-subject factor and familiarization (/gla.te/ versus /gla.tu/) as the between-subject factor. Results revealed a significant interaction of Parse Type x Familiarization, $F(1,30) = 6.297, p = .018$, while no other main effect was obtained (see Figure 3). As predicted, the group of infants familiarized with the affixed form /gla.te/ looked significantly longer while listening to the /glat/ stem trial than to the /gla/ partial stem trial, paired $t(15) = 2.724, p = .016$. For the group of infants familiarized with the unaffixed /gla.tu/ form, looking times for /gla/ and /glat/ were not significantly different, paired $t(15) = .913, p = .376$. Results are shown in Figure 3.

Results of Experiment 4 (Shi & Marquis, 2009) demonstrate that by 14 months of age, French-learning infants understand morphological alternations. They can link forms that are related in highly regular verb paradigms. These findings follow the decomposition theory of morphological learning that we argue for. The high token frequency of the suffix /e/ co-occurring with the high type frequency of verb stems that infants have encountered during their prior linguistic experience should have enabled them to parse the stem /glat/ and /e/ suffix, and to interpret the form /glat/ as being associated with the affixed form /gla.te/.

Infants who were familiarized with the unaffixed /gla.tu/ responded quite differently. They correctly treated /gla.tu/, /glat/ and /gla/ as separate, unrelated lexical items.

5. General discussion

This article concerns the most fundamental aspects of verb learning. We attempt to understand how infants might resolve the problems of segmentation and morphological
variations of verb forms. At the beginning stage of language learning, infants must determine the possible lexical units of their language from continuous speech, that is, to find the units that will allow them to build a vocabulary. Segmentation of even the most stable word forms, those that do not go through morphological changes, is itself a challenging task for infants, as the acoustic and phonetic realizations of words naturally vary due to many factors such as sentential position, co-articulation, phonological operations, focus, speech rate, affect, speaker differences, etc. The story is particularly complex for languages that involve a high degree of morphological variations. Inflectional morphology, for example, may pose a major challenge for early verb learning. Infants not only have to segment verb forms, but must also learn that certain variable forms are associated in verb paradigms (e.g., call, calls, calling, called).

In the first section of this article, we discussed two theoretical positions for verb representations, the non-decomposition theory and the decomposition theory. According to the non-decomposition theory, morphological alternations of the same verb (e.g., call, calls, calling, called) are encoded in the lexicon as non-separable single units. Therefore, the implication for acquisition is that infants would most likely need semantic evidence to learn that these forms are related in a verb paradigm. We believe that the reliance on semantics for early verb learning is unlikely since there is evidence that verb meanings are harder to induce than noun meanings (e.g., Gillette, Gleitman, Gleitman, & Lederer, 1999). Instead, we propose a theory of decomposition without semantics for initial verb acquisition. Specifically, we suggest that verbs that contain internal regular morphology (such as calling) are decomposed into stems and suffixes at the earliest stage of learning. The mechanism of this learning involves the frequency of stem and suffix distributions: high token frequency of the suffixes co-occurring with high type frequency of the stems leads to the word internal decomposition. We also suggest that infants should interpret the segmented stem (from affixed forms) and the bare stem of the same verb as variants of the same abstract unit, and
that the learning system would subsequently expect these variants to receive the same core meaning. Moreover, we assume continuity of the decomposed morphological representations from infancy to adulthood. That is, the decomposed units are represented in the lexicon during initial learning by infants and remain so in the mature lexicon.

The empirical evidence that we discussed in this article supports the decomposition theory. In our experiments, we examined the knowledge of verb morphology in infants at the age when they hardly have any meaningful vocabulary. We showed that infants who acquire French, a language with considerable verb morphological variations, are capable of parsing stems from affixed forms and associate such forms with the bare stem. They are able to do so on the basis of distributional analyses, without relying on verb meaning. Infants’ responses in our experiments reflected their prior learning. Thus, when hearing a novel verb ending with the regular verb suffix /e/ during our experiments, infants parsed them into the stem and suffix because they had months of accumulated exposure to this highly frequent suffix co-occurring with many different stems. In other words, the different stems that infants had encountered in their natural environment were of high type frequency, which lowered the transitional probability between any stem and the suffix, therefore supporting the decomposition of affixed verbs into smaller morphological units. Furthermore, we obtained direct evidence supporting the knowledge of the French verb paradigm involving /e/ in 14-month-old infants. After hearing a suffixed novel verb form and then being given the choices of a related bare stem versus an unrelated but partially matching word form, infants correctly linked the bare stem and the affixed form of the verb. Overall, the combined results from our perceptual experiments provide robust evidence for the decomposition account of infants’ initial learning of verb form alternations.

In addition to evidence from perceptual learning, children’s speech production also supports the decomposition theory. Infants’ early speech is known to be telegraphic, lacking
functional morphemes (e.g., Brown, 1973). Infants’ spontaneous speech typically omits inflections in cases where the usage is obligatory. In controlled speech production experiments, toddlers only dropped real grammatical suffixes but not nonsense suffixes (e.g., Gerken, Landau, & Remez, 1990), suggesting that they had separate representations of the stem and suffix, but they had a whole representation for disyllabic mono-morphemic forms. These observations with older infants are consistent with the findings from our perceptual experiments with preverbal and early verbal infants, indicating that children decompose stems and suffixes from the beginning of verb learning. Furthermore, omission patterns in production show that children associate stems from various contexts with the same word meaning. Additional evidence for the decomposition theory can be seen in children’s overgeneralization productions. For example, English children often produce incorrect forms such as *finded as the past tense of find (e.g., Pinker, 1995), indicating that regular suffixes are separately represented and serve as the default paradigms for English children. Forms of this kind are not produced by adults, and must therefore be the result of a rule-generating process in children. Similarly, French children often wrongly conjugate irregulars with regular endings as in *je l’ai batté instead of je l’ai battu ‘I’ve beaten him/her’ (e.g., Royle, 2007). These errors also indicate that verbs are decomposed into stem and affixes.

Decomposition appears to be a basic element for most morphological theories in linguistics and psycholinguistics. Many existing theories focus on adult representations, and therefore are not concerned with the question of decomposition versus non-decomposition of inflected words during acquisition. Nevertheless, they usually contain inflection rules, at least for regular verbs, and the rules would presumably operate with some kind of representations corresponding to suffixes and stems. This would be the case for the morphological framework under generative phonology (Chomsky & Halle, 1968), which posits rules for regular as well as irregular verbs. Stems and inflectional morphemes would logically be the representational
units for the rule constructs, knowledge that is most likely acquired during childhood. Unlike models under generative phonology, Bybee’s theory (Bybee, 1985, 1988) considers word frequency as a crucial factor that influences adult lexical representations. Low-frequency words with regular inflections are derived in the lexicon, but high-frequency words with regular inflections are stored as a whole. Logically, the high-frequency regulars were once infrequent at the early stage of acquisition. Thus, they may have been represented in decomposed units before the frequency of these affixed words reached the high-frequency threshold. But regardless of whether Bybee’s theory would regard these regulars as decomposed or not at the initial stage of acquisition, decomposition must have occurred at some stage that enabled the learning of morphological rules, which her theory uses for the derivation of low-frequency words with regular inflections. The rule-rote model by Pinker (e.g., Pinker, 1991; Pinker & Prince, 1994) can apply directly to early morphological acquisition: irregular verbs are rote memorized, and non-decomposed, in the lexicon, whereas regular verbs are generated by a rule process that must involve decomposed suffixes and stems. Our work explains an acquisition stage prior to the knowledge in Pinker’s model, we propose a learning mechanism by which infants establish these representational units and reach a basic understanding of morphological alternations.

In sum, our goal here is to develop a model of early morphological learning. We approach this acquisition problem by studying infants’ initial segmentation of verb forms and their interpretation of verb morphological alternations. Our empirical findings demonstrate that infants begin to parse verbs into decomposed stems and suffixes by 11 months of age, and they have rudimentary knowledge of regular verb paradigms by 14 months of age. We further show that this learning is entirely based on distributional analyses of the input without the need for semantics. The fact that infants in our studies treated morphologically alternating forms of a verb as variants of the same verb suggests that they should then expect these forms
to have the same core meaning, an outcome which we predict for future experiments that specifically test infants’ interpretation of word meaning. The remarkable morphological knowledge that we demonstrated in young infants constitutes an essential part of the early grammatical representations, which may directly impact the subsequent acquisition of semantics as well as more refined morphosyntactic structures.

Acknowledgment

This research was supported by grants from NSERC, SSHRC, FQRSC, and CFI to the second author, and by an FQRSC scholarship to the first author. We wish to thank Andréane Melançon, Elena Kulagina, Bruno Gauthier, and Mélanie Lepage for research assistance.
### Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1SG</td>
<td>first person singular</td>
</tr>
<tr>
<td>2PL</td>
<td>second person plural</td>
</tr>
<tr>
<td>2SG</td>
<td>second person singular</td>
</tr>
<tr>
<td>3SG</td>
<td>third person singular</td>
</tr>
<tr>
<td>AUX</td>
<td>auxiliary verb</td>
</tr>
<tr>
<td>CVC</td>
<td>consonant-vowel-consonant</td>
</tr>
<tr>
<td>CCVC</td>
<td>consonant- consonant-vowel-consonant</td>
</tr>
<tr>
<td>CCVCV</td>
<td>consonant- consonant-vowel-consonant-vowel</td>
</tr>
<tr>
<td>IMP</td>
<td>imperative</td>
</tr>
<tr>
<td>INF</td>
<td>infinitive</td>
</tr>
<tr>
<td>PP</td>
<td>past participle</td>
</tr>
<tr>
<td>s</td>
<td>seconds</td>
</tr>
</tbody>
</table>
References


DOI:10.1207/s15327078in1002_5.

Appendix 1. Experiment 1 stimuli

**Familiarization:** either /bif/ or /tar/

**Test:**

La jolie maman /bif/ les poèmes.
‘The pretty mommy cross out-3SG the poems.’

Elle /bif/ les verbes.
‘She cross out-3SG the verbs.’

Évidemment on /bif/ la virgule.
‘Obviously we cross out-3SG the comma.’

C’est le juron qu’il /bif/.
‘It’s the curse that he cross out-3SG.’

Le petit bébé /bif/ le graffiti.
‘The little baby cross out-3SG the graffiti.’

La syllabe je /bif/.
‘The syllable I cross out-1SG.’

/bif/-tu le calembour?
‘Do you cross out-2SG the pun?’

&

La nouvelle maman /tar/le magot.
‘The new mommy weigh-3SG the pile.’

Il /tar/ les grappes.
‘He weigh-3SG the clusters.’

Visiblement je /tar/ la parcelle.
‘Visibly I weigh-1SG the fragment.’

C’est le lot qu’elle /tar/.
‘It’s the share that she weigh-3SG.’

Le gentil bébé /tar/ les gondoles.
‘The nice baby weigh-3SG the gondola.’

Les rouets je /tar/.
‘The spinning wheels I weigh-1SG.’

/tar/-tu la coquille?
‘Do you weigh-2SG the shell?’
Appendix 2. Experiment 2 Condition 1 stimuli

**Familiarization:** either /glyt/ or /trid/

**Test:**

Maman a /gly.te/ le magot.
‘Mommy AUX /gly.te/NONCE.PP the pile.’
/gly.te/ les grappes est amusant.
‘/gly.te/NONCE.INF the grapes is amusing.’
T’as /gly.te/ la parcelle.
‘You AUX /gly.te/NONCE.PP the parcel.’
C’est le lot qu’on a /gly.te/.
‘It’s the lot that we AUX /gly.te/NONCE.PP.’
Bébé va /gly.te/ les gondoles.
‘Baby AUX /gly.te/NONCE.INF the gondolas.’
Les roues, j’ai /gly.te/.
‘The wheels, I AUX /gly.te/NONCE.PP.’
/gly.te/ la coquille!
‘/gly.te/NONCE.2PL.IMP the shell!’

&

Maman a /tri.de/ les poèmes.
‘Mommy AUX /tri.de/NONCE.PP the poems.’
/tri.de/ les verbes est amusant.
‘/tri.de/NONCE.INF the verbs is amusing.’
T’as /tri.de/ la virgule.
‘You AUX /tri.de/NONCE.PP the comma.’
C’est le jeu qu’on a /tri.de/.
‘It’s the game that we AUX /tri.de/NONCE.PP.’
Bébé va /tri.de/ le graphème.
‘Baby AUX /tri.de/NONCE.INF the grapheme.’
La phrase j’ai /tri.de/.
‘The sentence I AUX /tri.de/NONCE.PP.’
/tri.de/ le juron!
‘/tri.de/NONCE.2PL.IMP the curse word!’
Appendix 3. Experiment 2 Condition 2 stimuli

Familiarization: either /glyt/ or /trid/

Test: *Maman a /gly.tu/ le magot.
*Mommy AUX /gly.tu/NONCE.CCVCV the pile.’
*/gly.tu/ les grappes est amusant.
*The grapes is amusing.’
*T’as /gly.tu/ la parcelle.
*You AUX /gly.tu/NONCE.CCVCV the parcel.’
*C’est le lot qu’on a /gly.tu/.
*It’s the lot that we AUX /gly.tu/NONCE.CCVCV.’
*Bébé va /gly.tu/ les gondoles.
*Baby AUX /gly.tu/NONCE.CCVCV the gondolas.’
*Les roues, j’ai /gly.tu/.
*The wheels, I AUX /gly.tu/NONCE.CCVCV.’
*/gly.tu/ la coquille!
*The shell!’

&

*Maman a /tri.du/ les poèmes.
*Mommy AUX /tri.du/NONCE.CCVCV the poems.’
*/tri.du/ les verbes est amusant.
*The verbs is amusing.’
*T’as /tri.du/ la virgule.
*You AUX /tri.du/NONCE.CCVCV the comma.’
*C’est le jeu qu’on a /tri.du/.
*It’s the game that we AUX /tri.du/NONCE.CCVCV.’
*Bébé va /tri.du/ le graphème.
*Baby AUX /tri.du/NONCE.CCVCV the grapheme.’
*La phrase, j’ai /tri.du/.
*The sentence, I AUX /tri.du/NONCE.CCVCV.’
*/tri.du/ le juron!
*The curse word!’
Figure Captions

Figure 1. 11-month-old infants’ orientation times (means with standard errors of the means) to the test sentences containing the familiarized target verb and non-familiarized verb (Experiment 1).

Figure 2. 11-month-old infants’ orientation times (means with standard errors of the means) to the test sentences containing the familiarized target verb and non-familiarized verb (Experiment 2). The two left columns are the real suffix /e/ Condition 1, and the two right columns are the nonsense suffix /u/ Condition 2.

Figure 3. 14-month-old infants’ orientation times (means with standard errors of the means) to /glat/ and /gla/ (Experiment 4). The left two columns are the results for the group familiarized with the verb affixed with the real suffix /e/, and the right two columns are the results for the group familiarized with the disyllabic mono-morphemic form.
Figure 1.
Figure 2.

Experiment 2: 11-month-old infants' parsing of suffixed verbs

- Sentences containing the familiarized target root
- Sentences containing the non-familiarized root

Looking time (sec.)

<table>
<thead>
<tr>
<th>Looking Time (sec.)</th>
<th>/gly.te-tri.de/ vs. Familiarization</th>
<th>/gly.tu-tri.du/</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* * n.s.
Experiment 4: 14-month-old infant's processing of verb morphological variations

Looking time (sec.)

/gla.te/ vs. /gla.tu/

Familiarization

* n.s.

Figure 3.