J. Child Lang. **36** (2009), 557–596. © 2008 Cambridge University Press
doi:10.1017/S0305000908009124 Printed in the United Kingdom

Liaison acquisition, word segmentation and construction in French: a usage-based account

JEAN-PIERRE CHEVROT*
LIDILEM, Université Stendhal, Grenoble, France

CELINE DUGUA
LLL/CORAL, Université d'Orléans, France

AND

MICHEL FAYOL
LAPSCO, Université Blaise Pascal & CNRS, Clermont-Ferrand, France

(*Received* 19 *January* 2006. *Revised* 20 *September* 2007. *First published online* 24 *October* 2008)

ABSTRACT

In the linguistics field, liaison in French is interpreted as an indicator of interactions between the various levels of language organization. The current study examines the same issue while adopting a developmental perspective. Five experiments involving children aged two to six years provide evidence for a developmental scenario which interrelates a number of different issues: the acquisition of phonological alternations, the segmentation of new words, the long-term stabilization of the word form in the lexicon and the formation of item-based constructions. According to this scenario, children favour the presence of initial CV syllables when segmenting stored chunks of speech of the type word1-liaison-word2 (*les arbres* 'the trees' is segmented as /le/+/zarbr/). They cope with the variation of the liaison in the input by memorizing multiple exemplars of the same word2 (/zarbr/, /narbr/). They learn the correct relations between the word1s and the word2 exemplars through exposure to the well-formed sequence (*un*+/*narbr*/, *deux*+/zarbr/). They generalize the relation between a word1 and a class of word2 exemplars beginning with a specific liaison consonant by integrating

[*] We should like to thank Ann Peters for the initial impetus she gave to this work as well as Marie-Hélène Côté, Bernard Laks, Yves Charles Morin and Sophie Wauquier-Gravelines for the interest they have shown in it. We would also like to thank the referees and the associate editor for their revealing and constructive comments. Address for correspondence: Jean-Pierre Chevrot, Université Stendhal, BP 25, 38040, Grenoble cedex, France. e-mail: jpchevrot@wanadoo.fr

https://doi.org/10.1017/S0305000908009124 Published online by Cambridge University Press

this information into an item-based schema (e.g. *un*+/nX/, *deux*+/zX/). This model is based on the idea that the segmentation of new words and the development of syntactic schemas are two aspects of the same process.

INTRODUCTION

Liaison in French is one of the sandhi phenomena that are present in many languages. In the field of linguistics, liaison is a recurrent theme in the study of adult phonology. Its modelling has proved to be an unavoidable test for all of the many theories proposed: Schane's first approach (1968), based on the framework of rule-based phonology, Tranel's proposals (2000), conceived in the light of constraint-based theory or the recent model proposed by Bybee (2001), which is based on the idea of construction. In this latter conception, recurrent chunks of speech which contain a liaison are memorized as a whole and this mode of storage explains why the corresponding liaisons are used so frequently and why they have been preserved in French.

Moreover, the variable nature of certain liaison consonants has been one of the preferred objects of research for corpus-based studies of phonological variation in adults (De Jong, 1994; Encrevé, 1988; Moisset, 2000). We therefore possess a good understanding of liaison in adults. In contrast, the question of its acquisition in children has long been neglected, despite the fact that this would provide us with a new paradigm for understanding the interplay between several developmental issues: the acquisition of phonological alternations, the segmentation of new words, the long-term stabilization of the word form in the lexicon and the formation of early syntactic constructions.

To help gain a better understanding of the specific constraints that liaison exerts on these developmental issues, we shall start by presenting its functioning in adult speech.

In French-speaking adults, liaison consonants appear between two words in connected speech. A necessary condition is that the right-hand word (hereafter word2) starts with a vowel when spoken in isolation. In contrast, the liaison consonant is never produced at the end of the left-hand word (hereafter word1) when situated at the end of an utterance or when it precedes a word2 that starts with a consonant. Similarly, this consonant is never pronounced at the start of word2 when it is located at the beginning of an utterance. For example, the liaison /t/ is pronounced between *petit* and *écureuil* in the sequence *petit écureuil* ([ptitekyʀœj] 'small squirrel'). However, in adult speakers, this /t/ is not pronounced in either *petit veau* ([ptivo] 'small calf') or in *il est petit* ([ilepti] 'he is small') or at the start of

https://doi.org/10.1017/S0305000908009124 Published online by Cambridge University Press

Ecureuil! ([ekyʀœj] 'Squirrel!'). When this liaison consonant (hereafter LC) is produced, it generally forms a syllable with the initial vowel of the following word: for instance, the syllabification of the sequence *petit écureuil* is [pti.te.ky.ʀœj]. It is important to emphasize that this syllabic linkage operates over the word boundary and means that the syllabic boundary no longer corresponds to the lexical boundary. Finally, both the possibility of producing a liaison and its phonetic nature (/n/, /z/ and /t/ in 99·7% of cases, Boë & Tubach, 1992) depend on the word1 'as if liaison belonged to it' (Tranel, 2000; our translation). For example, the word1s *un* 'a/one' or *aucun* 'none' both trigger a liaison in /n/, the word1s *petit* or *grand* a liaison in /t/, the word1s *gros* or *deux* a liaison in /z/, whereas *joli* or *beau* in the singular do not trigger any liaison when appearing as a word1. Even though these phonological facts have long been well known to linguists, the acquisition of liaison has never been studied in depth.

We start by presenting a developmental scenario which adopts a usage-based viewpoint (Kemmer & Barlow, 2000; Tomasello, 2003). According to the usage-based point of view, the generalizations formulated by children are thought to emerge from the accumulation, the analogous processing and the reorganization of concrete linguistic material consisting of memorized words or word sequences. We support the idea that children segment word sequences containing an LC by attaching this consonant to the start of various lexical exemplars of the word that follows in the speech stream (e.g. /nurs/, /zurs/ for the word *ours* 'bear'). As a result of prolonged contact with well-formed sequences encountered in the input, children generalize the link between a specific word1 and appropriate word2 exemplars (e.g. *un*+/nurs/, *deux*+/zurs/) before integrating this information into item-based constructions based on the word1 (e.g. *un*+/nX/, *deux*+/zX/). This model postulates that the segmentation of new words and the development of syntactic schemas are two aspects of the same process (Bates & Goodman, 1997; Tomasello, 2003: 93).

Our contribution will consist, first of all, of setting out the assumptions which underpin our usage-based approach. We shall then explain the proposed scenario with regard to the case of liaisons which follow determiners within nominal phrases. Finally, we shall present five experiments which test the central hypotheses of this scenario.

MAIN ASSUMPTIONS

Usage-based theories appear to act as a wide-ranging alternative to the idea of innateness. They postulate that language use is the key to understanding language organization. Of the principles common to the various strands of these theories, we shall present those which directly support our developmental model.

https://doi.org/10.1017/S0305000908009124 Published online by Cambridge University Press

One key postulate is that usage events, that is to say the specific utterances that speakers produce and hear, constitute the experience from which they construct their linguistic knowledge (Kemmer & Barlow, 2000). These usage events are concrete and specific in nature in the sense that they are formed from lexical units which possess a phonological content. On contact with them, children are thought to memorize concrete pieces of language formed from a sound sequence which is associated with a communicative intention (Tomasello, 2003). Depending on the specific case, these concrete pieces of language may correspond to a single adult word or, as has also been documented (Pine & Lieven, 1993), to a sequence of words memorized as a frozen expression. Pine & Lieven (1997) have noted that English-speaking children aged two to three years rarely use the nouns they know with the determiners *a* and *the* but instead seem to associate each noun with a specific determiner. Certain determiner–noun sequences could therefore function as frozen expressions, and the French language might amplify this phenomenon since, in this language, common nouns are almost always preceded by a determiner. In an analysis of the data collected during a diary study, Dugua (2006: 240) noted the emergence, at two to three years of age, of errors of the type *un la trompe* 'a the trunk' instead of *une trompe* 'a trunk' and *un la lampe* 'a the lamp' instead of *une lampe* 'a lamp'. These can be considered as the reuse of the memorized sequences *la-trompe* or *la-lampe* following the indefinite article *un*.

A second postulate relates to the early construction of linguistic knowledge. By reorganizing these concrete pieces of language, children become capable of producing new sequences. In doing so, they may combine multiple chunks under the same intonational contour or generalize more abstract schemas (Tomasello, 2003). During the schematization process, a relationship would be established between the memorized chunks on the basis of their phonological or functional similarities. For example, by linking together the sequences which share the French singular indefinite article (*un garçon*, 'a boy', *un livre* 'a book', *un ours* 'a bear', etc.), children are thought to identify a schema based on the lexical item *un*. This schema can be represented in the form *un*+X, where X is a slot in which new elements can be inserted. Since the schemas have no existence independently of the sequences from which they emerge (Kemmer & Barlow, 2000), this notation provides a convenient and concise way of describing the pattern which connects the memorized traces of language experiences.

A third postulate relates to the key status of the idea of construction, which is thought to be the basic unit of the linguistic system. Constructions are conventionalized patterns which establish a relation between form, meaning and function (Goldberg, 2003). The different types of constructions can be described by situating them on a continuum involving two

https://doi.org/10.1017/S0305000908009124 Published online by Cambridge University Press

dimensions: from the concrete to the abstract and from the simple to the complex (Croft & Cruse, 2004; Tomasello, 2003). They may therefore take the form of a word (simple and concrete), a sentence structure formed from abstract categories (complex and abstract), a fixed expression (complex and concrete) or a mixed configuration which combines phonologically specified elements with abstract categories (as in the case of item-based schemas, e.g. *un*+X). During the course of development, the constructions are organized into networks and become more complex and abstract.

These three postulates have two direct consequences. First, since the linguistic system is guided by the accumulation of language experiences, the frequency of the constructions is an important factor structuring the way it operates (Kemmer & Barlow, 2000). Second, since constructions organize information coming from multiple linguistic levels (phonology, lexicon, morphology, syntax), they involve an interactive conception of language development and functioning. In the developmental scenario which follows, we argue that the acquisition of prenominal liaisons reveals interactions between various linguistic levels: the learning of phonological alternations, the segmentation and stabilization of new words, and the formation of schemas based on determiners which prefigure the structure of the nominal phrase.

A TWO-STAGE DEVELOPMENTAL SCENARIO

One important task which young children have to perform is to construct, on the basis of the speech they hear in their environment, the form, function and meaning of the units in their lexicon. In some cases, the stress and distribution patterns as well as the pragmatic and referential salience of certain words would allow children to extract certain units directly from the utterances they hear in the surrounding environment. In other cases, children would memorize and process, as a single lexical entity, sequences which correspond to multiple words for adults. They then have to segment these sequences in order to extract the units which can be used in other utterances (Peters, 1985; Tomasello, 2003).

Little is known about the respective importance of the direct extraction process and the follow-up processing of memorized sequences and we shall accept that the early lexicon is made up of segmented units and chunks. It is precisely the segmentation of the chunks which contain a liaison that forms the object of the scenario that we propose here. In effect, children who are confronted with the French language have to memorize sequences such as *un-âne* ([œ̃nɑn] 'a donkey') which contain an LC (/n/ in this case). On the basis of these sequences, they have to extract a determiner and a noun which can be re-used in other utterances (Wauquier-Gravelines & Braud, 2005).

https://doi.org/10.1017/S0305000908009124 Published online by Cambridge University Press

Stage 1 – Segmenting exemplars of word2 and constructing general item-based schemas

During an initial developmental stage, the alternation of the different LCs and the dissociation between the lexical and syllabic boundaries which this induces would lead young children to segment various lexical exemplars of each word2.

In effect, the LC forms a consonant–vowel (CV) syllable with the start of word2 and this sequencing results in dissociation between the lexical boundary and the syllabic boundary. Thus *les arbres* 'the trees' is syllabified as [le.zarbr], with the syllabic boundary located in front of the /z/ and the lexical boundary after the /z/. We also know that the lexical segmentation process in babies makes use of the probable correspondence between syllabic boundaries and word boundaries (Mattys & Jusczyk, 2001). Moreover, consonants are more likely segmentation points than the vowels. Using the database Lexique (New, Pallier, Ferrand & Matos, 2001), one can establish that French nouns starting with a consonant are about three times more numerous than those starting with a vowel. For syllabic and statistical reasons, children would place a lexical boundary in front of the LC which, as a result, would be attached to the start of the lexical representation of the word2. It is therefore to be expected that children will retrieve the lexical form /zarbr/ after hearing a sequence such as *les arbres*.

However, the second characteristic of liaison – its variation as a function of word1 – also constrains the segmentation process. Children encounter each word2 preceded by different LCs, for example *arbre* 'tree' preceded by /n/ in *un arbre* 'a tree', by /z/ in *les arbres* 'the trees' and by /t/ in *petit arbre* 'small tree'. If they segment these sequences by applying a syllabic strategy in all cases (/œ̃.narbr/, /le.zarbr/, /pti.tarbr/), they will obtain multiple exemplars of each lexical unit: /narbr/, /zarbr/, /tarbr/. These consonant-initial exemplars may then be joined by the vowel-initial variant (/arbr/) of certain word2s which are heard in isolation or in contexts where they are preceded by a word1 which does not cause any liaison (*joli arbre* 'pretty tree').[1]

This segmentation of the word2s is correlative with the segmentation of the word1s. In the case of prenominal liaisons, the latter belong to the restricted class of the determiners. If segmentation results from the establishment of a relationship between memorized chunks which share certain phonological and semantic content, then we would expect these schemas to be formed on the basis of those words which are more likely to be present in a large number of chunks. The determiners could therefore become a concrete element in a schema of the form *un*+X (or *les*+X, *des*+X, etc.)

[1] For a *similar* scenario outline, based on the same idea of alternating variants of the word2s, see Morin (2003 [1998]).

https://doi.org/10.1017/S0305000908009124 Published online by Cambridge University Press

which results from the connection between chunks which contain the word *un* (*un garçon*, 'a boy', *un livre* 'a book', *un ours* 'a bear', etc.). These schemas reveal children's newly acquired ability to combine a word1 and a word2. In effect, they provide a slot X in which children can insert the segmented variants of the word2s. In the case of prenominal liaisons, they prefigure the internal structure of the NP and would seem to belong to the general class of item-based constructions which are characteristic of early syntactic development (Tomasello, 2003). At this first stage, these schemas would not contain any information about the liaison.

Stage 2 – Constructing schemas which generalize the relation between word1 and word2 exemplars

While children form general schemas and segment word2 variants, they continue to memorize frequent determiner+noun sequences, some of which contain a liaison: e.g. *les ours*, *les ânes*, *les amis* (/lezurs/, /lezan/, /lezami/) 'the bears', 'the donkeys', 'the friends'. By establishing a relation between these memorized sequences, they generalize a schema of the type *les*+/zX/, which specifies that the slot following the determiner *les* should be filled with a variant starting with /z/. This schema would make it possible to produce correct liaisons without having to encounter and memorize all the possible correct combinations of word1 and word2 variants.

The second stage would then be characterized by a more abstract structure which generalizes the relation between a specific word1 and a specific class of word2 exemplars (*les*+/zX/, i.e. *les* should be followed by an exemplar starting with a /z/, and *un*+/nX/, i.e. the determiner *un* should be followed by an exemplar starting with an /n/, etc.).

If we accept that the specific constructions which bear liaisons are in competition with more general early acquired schemas of the type *un*+X or *les*+X which contain no information about liaison, this developmental scenario is able to account for the four main types of liaison errors identified in case studies (Chevrot & Fayol, 2001; Dugua, 2006; Grégoire, 1947; Wauquier-Gravelines, 2003). Overall, the existence of these errors provides evidence that young speakers do not only reproduce memorized word1–word2 sequences but that they also engage in a segmentation and generalization process.

First, the insertion of the exemplar /narbr/ in the schema *les*+X results in the classic error of substitution [lenarbr]: *les arbres* 'the trees' is pronounced with the liaison /n/ instead of the expected liaison /z/. Second, the insertion of the vowel-initial exemplar /arbr/ in the schema *les*+X produces the omission error [learbr]: *les arbres* is pronounced without any liaison. Third, the insertion of the exemplar /narbr/ in the schema *le*+X results in the adjunction error [lənarbr]: *l'arbre* 'the tree' is pronounced with the LC

https://doi.org/10.1017/S0305000908009124 Published online by Cambridge University Press

/n/ after the singular definite article *le* which does not prompt any liaison in adults. Fourth, this developmental scenario accounts for another type of error: substitutions of initial consonants. These errors appear in word1–word2 sequences where the word1 induces a liaison and the word2 starts with a /z/ or an /n/ in adult speakers (e.g. *zèbre* 'zebra' or *nuage* 'cloud'). In these errors, the /n/ or the /z/ is replaced by a consonant which is compatible with the word1. For example, children may produce [œ̃nɛbr] for the sequence *un zèbre* 'a zebra' or [lezyaʒ] for the sequence *les nuages* 'the clouds'. They thus replace the initial /z/ of *zèbre* with the consonant /n/ which, in other contexts, is the LC induced by *un*, or the initial /n/ of *nuage* with /z/ which is the LC induced by *les*.[2]

These errors would result from the assimilation of the form *nuage* into the schema *les*+/zX/ or the form *zèbre* into the schema *un*+/nX/. They can be thought of as indices of the productivity of the constructions which carry information about liaisons. In order to produce the target sequence *les ours* ('the bears'), the child has to insert the exemplar /zurs/ in the schema *les*+/zX/. However, it is possible that the exemplar /zurs/ is not available in her lexicon, for example because the child usually hears this noun after a singular determiner, inducing an /n/ liaison, not a /z/. In this case, we assume that the child would create the missing exemplar /zurs/ by analogy, on the basis of the complete sets of exemplars she has already stored. The created exemplar is then inserted into the schema *les*+/zX/ with the correct sequence [lezurs] as the end product. The process is exactly the same when errors are produced. In order to produce the target sequence *les nuages*, the child creates the form [zyaʒ] from the stored form [nyaʒ], thus resulting in the error [lezyaʒ].

ALTERNATIVE VIEW AND FREQUENCY EFFECT

Wauquier-Gravelines & Braud (2005) have proposed an alternative concept within the framework of autosegmental phonology. This idea holds that liaison errors are one of the first manifestations indicating that prosodic positions are independent of the contents of the segments and thus prefigure the future status of the liaison which is conceived of as a floating consonant in the autosegmental framework. In line with the principle that syllables with an onset are preferred, children would favour the presence of a noun-initial consonant when determining the boundary between the determiner and the noun. A prosodic position C corresponding to the liaison would

[2] The specific schemas of the type *les*+/zX/ predict that all the initial consonants of the noun (and not only /n/, /z/, /t/ and /l/) could be replaced with a liaison consonant. Indeed, Dugua (2006: 136) observes these errors in the everyday speech of a child: [dɑ̃lezɛstɔʀɑ̃] (*dans les restaurants*, 'in the restaurants', 3;2), [lezɑg] (*les vagues*, 'the waves', 3;7).

https://doi.org/10.1017/S0305000908009124 Published online by Cambridge University Press

then be associated with the syllabic onset of the noun. Children would then use various means to assign a concrete content to this abstract position, thus leading to liaison variations and errors: harmony with a noun-internal consonant (*les éléphants* 'the elephants' produced as [lefefefɑ̃]), use of a default yod, or probabilistic use of the most frequent liaison. Later, the influence of morphology would trigger the learning of the correct liaisons. These would be encoded in the form of floating consonants at the end of the determiners and adjectives when children discover 'that the determiner *un* takes a nasalized consonant in the feminine, that adjectives such as *petit*, *grand*, *gros* take a consonant of inflection … during derivation operations' (Wauquier-Gravelines & Braud, 2005; our translation). For example, children would encode a /t/ at the end of *petit* 'small' when they become aware that the feminine form *petite* contains the consonant /t/.

Chevrot, Chabanal & Dugua (2007) have empirically tested two differences which oppose this autosegmental conception with the exemplar-based conception. In the exemplar-based model, the MULTIPLE lexical representation associated with the noun *ours* is the combination of a number of CONCRETE sequences – /urs/, /nurs/, /zurs/, /turs/. The phonetic material corresponding to the liaison consonants specifically belongs to these exemplars. In the autosegmental conception, children assign a phonetic content to an abstract initial position either on the basis of context or by default. However, this phonetic content is not related to the lexical representation of the noun and does not belong to it. Chevrot *et al.* (2007) have shown that there is a relation between the phonetic nature of children's errors and the frequency of the liaisons which precede a given noun in the input. Using adults' intuitive judgments, these authors established that certain nouns occur more frequently in the plural (e.g. *arbre* 'tree'). In the input which they hear, children more frequently encounter these 'plural-oriented' nouns after the liaison /z/ which is typical of determiners which indicate plurality (*les*, *des*, 'the', *deux*, 'two', etc.). The analysis of the errors produced by a little girl from the age of two to six years reveals a high level of correlation between the 'plural-orientation' of the nouns and the probability that they will be preceded by an incorrect /z/ liaison in singular contexts (e.g. *un arbre* 'a tree' produced as [œ̃zɑrbr]). Chevrot *et al.* (2007) consider that this result is more in line with an exemplar-based than with an autosegmental account. In effect, if the phonetic content of the liaisons depends on the following noun, it is difficult to imagine that it starts with an abstract position with no phonetic content. Additional data supporting the exemplar-based view will be provided in the present paper.

The second major difference between the two models relates to the way the correct liaisons are initially learned. In the exemplar-based models, this learning results from the memorization of sequences consisting of a word1 and the appropriate word2 exemplar (e.g. *un*+/nurs/, *deux*+/zurs/).

https://doi.org/10.1017/S0305000908009124 Published online by Cambridge University Press

Children should therefore produce more correct liaisons in word1–word2 sequences which they hear and use frequently. In the autosegmental model, children encode a floating consonant at the end of word1 thanks to the morphological relations between this word and its derivative or inflected forms. This process depends on the existence of derived or inflected forms of word1 but not on the frequency of word1–word2 sequences. Analyzing the speech of a small girl aged between three and four years, Chevrot *et al.* (2007) observed the two main contexts in which liaisons are obligatory: after determiners and after clitics (e.g. the /n/ liaison in *on a soif* 'we are thirsty'). The authors noted a high level of correlation between the number of correct liaisons following these units and the distributional possibilities to the right of them. For example, the highest level of correct liaisons (97%) was observed after the clitic *en*, which is the unit with the most restricted number of possibilities to the right of it: *en* appears 104 times in the corpus and is followed by only 7 different word2s. Even though this pronoun has no derived form and is never inflected, children have no trouble using the /n/ which it activates. In comparison, the possible combinations to the right of the determiner *un* are less restricted: it appears 49 times in the corpus and precedes 12 word2s. Despite the fact that *un* is associated with the inflected form of the feminine *une*, the level of correct /n/ liaisons after the determiner (82%) is lower than the level after *en*. Fougeron, Goldman & Frauenfelder (2001) observed a similar result in the speech of ten adults: liaison occurs more often in frequent word1–word2 sequences.

Another frequency effect established by studies yielding convergent results is that the /n/ liaison is more prominent than the /z/ and /t/ liaisons in the errors made by children (Chevrot & Fayol, 2001; Dugua, 2006). As Chevrot & Fayol (2001) have noted, more obligatory /n/ and /z/ liaisons than /t/ liaisons are correctly realized in colloquial adult speech. However, /z/ appears later than /n/ in the phonological inventory of French children (Vinter, 2001). The early availability of /n/ would therefore appear to result from the interaction of frequency in the input – which penalizes /t/ – and the order of acquisition – which penalizes /z/. An additional factor in favour of /n/ could be that this consonant is a more likely segmentation point than /z/: whatever the vowel V, more French words start with /nV/ than with /zV/ (for the statistical details, see Chevrot & Fayol, 2001).

Overall, these earlier results point out that frequency is an important issue in the acquisition of liaison. Given that a liaison context occurs every sixteen words in adult speech (Boë & Tubach, 1992), the phenomenon is sufficiently frequent for children to identify regular patterns in its use. However, these influences vary as a function of the context. For instance, /z/ errors are nearly as prevalent as /n/ errors in front of plural-oriented nouns. Consequently, in order to identify these influences more precisely, we need a fine-grained description of the input frequency of the three main

https://doi.org/10.1017/S0305000908009124 Published online by Cambridge University Press

LCs in different contexts, especially in the case of obligatory liaisons on which the usage-based scenario focuses and preferably in child-directed speech. Unfortunately, the numerous corpus studies available in the liaison field do not fulfil these conditions. The majority of them focus on optional liaisons (Encrevé, 1988; Moisset, 2000) in the formal speech of educated speakers (Encrevé, 1988; Malécot, 1975). Moreover, the authors of these studies rarely publish tables to indicate the frequency of each LC in precise contexts and none of them analyzes liaisons in the speech directed to the children. Thus, there is no available corpus that provides sufficiently precise information concerning frequency of the obligatory liaisons in the input.

In the sections below, we present five experiments which complement earlier studies of liaison acquisition. Unlike in the preliminary work of Chevrot *et al*. (2007), the main goal of these experiments was not to address the frequency issue. Instead, the purpose of this paper is to provide more convincing data in support of the usage-based scenario.

The first experiment was a priming experiment involving thirty children aged three to four years. The aim of this experiment was to establish that liaisons are encoded at the start of the word2 exemplars in alternation in the early lexicon. The second experiment involved the use of imperatives. Two hundred children aged between two and six years had to call animal puppets whose names begin with a vowel in adults (*Ours, viens ici!* 'Bear, come here!'). This task induces the production of word2 not preceded by a word1. The aim was to identify the development of the availability of consonant-initial variants (*[nurs]*, *viens ici!*) and vowel-initial variants (*[urs]*, *viens ici!*). The third experiment took the form of a pseudonoun segmentation task involving determiner–noun liaison contexts. It was performed by the same sample of 200 subjects. It tested the hypothesis that constructions of the type *un*+X, which include no information about the LC, emerge earlier than constructions of type *un*+/nX/, which include this type of information. The fourth experiment involved the same sample and indicated changes in errors consisting of the substitution of fixed initial consonants (*un zèbre* 'a zebra' pronounced [œ̃nɛbr]). If these errors result from the generalization of constructions of the type *un*+/nX/, their developmental profile should indicate a regularization process with an increase in occurrences followed by a decline. The fifth experiment explored the relation between the production of generalization errors and judgments of acceptability in 123 children aged between two and six years. Its aim was to check whether children who produce [œ̃nɛbr] know that the word *zèbre* starts with a /z/.

All the experiments were conducted in a quiet room in nursery schools. The first four experiments were conducted by one of the authors and the fifth by linguistics students. The same sample of 200 children took part in experiments 2, 3 and 4. They were administered in a single session but

https://doi.org/10.1017/S0305000908009124 Published online by Cambridge University Press

separated by relaxation breaks. In all cases, the children were recorded individually and their responses were subsequently transcribed by adults.

EXPERIMENT 1: ERROR ELICITATION IN THREE PRIMING CONDITIONS

Almost all the generative phonological models proposed over the last forty years have assumed that the LC is attached to word1 in adult speakers (Schane, 1968). Some of them support the idea of lexically independent LCs (Côté, 2005). A smaller number of models assume that the LC is lexically attached to word2 (Ternes, 1977). Our usage-based conception of children's liaisons holds that various LCs are attached to the onset of various exemplars of the word2. In contrast, the autosegmental phonology framework claims that an abstract consonant position, located at the onset of the single lexical representation of the word2, is filled with a variable phonetic content retrieved from the context (Wauquier-Gravelines & Braud, 2005). An error elicitation experiment performed with three priming conditions should make it possible to determine which of these possibilities best reflects the liaison errors of children aged three to four years.

Task and predictions

The error elicitation task consisted of a primed picture naming task. For half of the items, the children produced a noun preceded by the determiner *un* 'a' after hearing the same or another noun preceded by the determiner *deux* 'two'. For the other half of the items, they heard the sequence *un*-noun and produced the sequence *deux*-noun. The task involved three priming conditions (see Table 1) which we illustrate here using the example of the target *un ours* 'a bear', for which the expected liaison is /n/ and the expected LC substitution error takes the form of the replacement of /n/ by /z/ ([œ̃zuʀs]). In the lexico-phonological condition, one and the same vowel-initial noun followed the determiner in both the prime and the target. The children heard *deux ours* 'two bears' with a /z/ liaison and then had to produce *un ours*. In the phonological condition, the noun in the prime differed from the noun in the target. The children heard *deux arbres* 'two trees' with a /z/ liaison and then had to produce *un ours*. In the control condition, the children produced *un ours* 'a bear' after hearing a determiner–noun sequence which did not contain any LC (a sequence in which the noun starts with a consonant, e.g. *deux balais* ([døbale] 'two brooms').

If we accept that the presence of the intruding unit in the prime favours its activation and causes the expected error, the three possible lexical statuses of the LC (independence, word1, word2) predict three error patterns which are listed in Table 1. Only if the LC is attached to the start

https://doi.org/10.1017/S0305000908009124 Published online by Cambridge University Press

https://doi.org/10.1017/S0305000908009124 Published online by Cambridge University Press

TABLE 1. *Experiment 1 – Presence of the intruding unit in the prime as a function of the three possible lexical statuses of the LC (target: un ours 'a bear' with LC /n/ ([œ̃nuʀs]); expected error: replacement of /n/ by /z/ ([œ̃zuʀs])*

	Lexico–phonological condition heard prime: [døzuʀs] (*deux ours* 'two bears')		Phonological condition heard prime: [døzaʀbʀ] (*deux arbres* 'two trees')		Control condition heard prime: [døbale] (*deux balais* 'two brooms)	
Replacement unit in the error as a function of the locus of lexical attachment of the LC	Units activated by the prime	Replacement unit present in the prime	Units activated by the prime	Replacement unit present in the prime	Units activated by the prime	Replacement unit present in the prime
Model 1: LC at beginning of word2 variants	[dø]	yes	[dø]	no	[dø]	no
Replacement unit: /zuʀs/ (zours)	[zuʀs]		[zaʀbʀ]		[bale]	
Model 2: LC independent	[dø]	yes	[dø]	yes	[dø]	no
Replacement unit: /z/	[z]		[z]		[bale]	
	[uʀs]		[aʀbʀ]			
Model 3: LC at end of word1 variants	[døz]	no	[døz]	no	[dø]	no
Replacement unit: /œ̃z/ (un-z)	[uʀs]		[aʀbʀ]		[bale]	

of word2 (model 1 in Table 1) would we expect the number of errors to increase in the lexico–phonological condition, and only in this condition. In effect, if the LC is attached to word2, the intruding lexical unit that causes the error [œ̃zuʀs] is /zuʀs/. This unit is present in the prime in the lexico–phonological condition (*deux ours* [døzuʀs]) but not in the phonological condition (*deux arbres* [døzaʀbʀ]). In contrast, if the LC is attached neither to the lexical representation of word1 nor to that of word2 (independent LC, model 2 in Table 1), then the intruding unit that causes the error [œ̃zuʀs] is restricted to the phoneme /z/. Since the unit /z/ is present in the prime in both the lexico–phonological condition (*deux ours* [døzuʀs]) and in the phonological condition (*deux arbres* [døzaʀbʀ]), there should be more errors in these two conditions than in the control condition. Finally, if the LC is attached to the end of word1 (model 3, Table 1), then the intruding unit that causes the error [œ̃zuʀs] is /œ̃z/. Since this unit is not present either in the lexico–phonological condition or in the phonological condition, priming should not modify the frequency of errors compared to the control condition.

Subjects and procedure

The participants consisted of thirty monolingual French-speaking children, aged 3;1 to 4;6 ($M = 3;8$).

The task consisted of producing thirty-two word1–word2 sequences (determiner–masculine noun) on the basis of sixteen pictures, eight of which represented the eight word2s drawn once and the other eight the same word2s drawn twice. Four of the word2s started with a vowel, with two being monosyllabic and two trisyllabic (*ours* 'bear', *arbre* 'tree', *écureuil* 'squirrel', *éléphant* 'elephant'), and four with a consonant (*singe* 'monkey', *cochon* 'pig', *balai* 'broom', *ballon* 'ball'). The drawings representing the four word2s that started with a vowel had already been used in picture naming tasks involving the production of liaisons. In an analysis of the data collected by Dugua (2002) from 200 children aged from three to six years, Nardy (2003) established that the non-response level corresponding to each of these four words varies between 5% and 13%. By using a limited number of words that could be represented by easily recognized drawings, we could be sure that the naming task would be accomplished successfully. The Appendix displays certain characteristics of the target word1s (frequency) and word2s (frequency and ratio of plural-orientation for the vowel-initial nouns) used in experiment 1 as well as in the following ones. Although the frequencies are not homogeneous in the complete set of stimuli, it should be noted that, in experiment 1, as well as in the other ones, we never oppose two experimental conditions or two age groups on different sets of stimuli.

https://doi.org/10.1017/S0305000908009124 Published online by Cambridge University Press

In each priming condition, the children twice produced each combination of each of the eight word2s with the two word1s, *un* 'a' and *deux* 'two'. As the consonant-initial word2s alternated with the vowel-initial word2s, two LCs were never produced in succession.

In the lexico–phonological condition, the experimenter drew, for example, a picture of a single bear and spoke the prime: e.g. *Sur cette image, il n'y a pas deux ours* (with the correct /z/ liaison) *mais* … 'In this picture there are not two bears, but …'. The child then had to produce the target sequence *un ours* 'one bear' with the determiner *un* that requires the LC /n/ after hearing *deux ours*, which has a /z/ liaison, in the sequence [døzurs]. The experimenter then presented a picture representing a single exemplar of one of the word2s starting with a consonant, for example *cochon* 'pig', saying: *Sur cette image, il n'y a pas deux cochons* (without LC) *mais* … 'In this picture there are not two pigs, but …'. The child then produced the sequence *un cochon* 'a pig', which does not require an LC in adults. This series of vowel-initial words inducing a liaison and consonant-initial words inhibiting a liaison continued until each of the four vowel-initial target words (*ours* 'bear', *arbre* 'tree', *écureuil* 'squirrel', *éléphant* 'elephant') had been produced twice with the determiner *un*. The same operation was then repeated with the plural/singular order inverted, i.e. the prime was heard with the determiner *un* and the target had to be produced with the determiner *deux*. The order of presentation of the vowel-initial and consonant-initial items was randomized in both of these number conditions (plural heard/singular produced and vice versa).

In the phonological condition, the procedure was the same but for one difference: the heard noun and the noun to be produced were not the same. For example, in the case of the vowel-initial nouns, the children heard the sequence *deux ours* 'two bears' with a /z/ liaison and had to produce the sequence *un arbre* 'a tree' which requires an /n/ liaison. The same was true of the consonant-initial nouns: for example, the children heard *deux cochons* 'two pigs' and had to produce *un balai* 'a broom'. As in the preceding condition, the item presentation order was randomized. However, the correspondence between the heard noun and the produced noun was not. As far as the vowel-initial words are concerned, the noun *arbre* could only be produced after hearing *ours* and vice versa. Similarly, the noun *éléphant* could only be produced after hearing *écureuil* and vice versa.

The same procedure was employed for a third time in the control condition. This time, the participants produced sequences consisting of a determiner and a vowel-initial word that induced a liaison (e.g. *un ours*) after hearing a sequence consisting of a determiner and a consonant-initial word that did not induce a liaison (e.g. *deux balais*). Thus, the children had to produce a liaison without previously hearing another liaison. Sequences in which the target to be produced did not contain a liaison (*un ballon* heard,

https://doi.org/10.1017/S0305000908009124 Published online by Cambridge University Press

TABLE 2. *Experiment 1 – Error elicitation in three priming conditions (mean scores, standard deviations)*

	/n/ replaces /z/ errors	/z/ replaces /n/ errors	Total
Lexico–phonological condition	2·17 (2·62)	1·17 (1·91)	3·34 (3·42)
Phonological condition	1·67 (2·26)	0·40 (1·04)	2·07 (2·46)
Control condition	1·67 (2·29)	0·33 (0·66)	2·00 (2·30)
Total	5·50 (6·91)	1·90 (3·19)	

un singe produced) alternated with sequences in which the target required a liaison (*un ballon* heard, *un arbre* produced). In the same way as in the first two conditions, the item presentation order was randomized, whereas the correspondence between the produced noun and the heard noun was not. As far as the vowel-initial nouns are concerned, the produced nouns *arbre, ours*, *écureuil, éléphant* were always preceded by the heard nouns *cochon, ballon*, *balai, singe* respectively.

Overall, each child therefore produced the same thirty-two sequences in each priming condition: sixteen with a consonant-initial noun (e.g. *un cochon*) and sixteen with a vowel-initial noun (e.g. *un ours*). Only the latter, i.e. those requiring a liaison, were processed. Of these, eight responses in each condition were associated with a target liaison /z/ following the determiner *deux* and the other eight with a liaison /n/ following the determiner *un*.

Half of the children were presented with the lexico–phonological condition followed by the phonological condition while the other half were presented with the phonological condition followed by the lexico–phonological condition. All the children were then subjected to the control condition. In each condition, the plural/singular order of the targets was counterbalanced. The three conditions were presented in a single session that lasted about twenty minutes.

RESULTS

When all three conditions were confounded, the thirty subjects produced 222 expected errors, i.e. 15% of the word1–word2 sequences in which word2 started with a vowel. The mean number of errors per subject and per condition was 2·43 for sixteen sequences produced.

When the target liaison was an /n/ liaison after the determiner *un*, the expected errors were of the type /z/ replaces /n/. In contrast, when the target was a /z/ liaison after the determiner *deux*, the expected errors were of the type /n/ replaces /z/. The scores on expected errors (Table 2) were analyzed by means of a 3 priming conditions (control, lexico–phonological,

https://doi.org/10.1017/S0305000908009124 Published online by Cambridge University Press

phonological) × 2 target liaisons (/n/ after the determiner *un* vs /z/ after the determiner *deux*) ANOVA. We report both by-participants (F1) and by-items (F2) values.

Errors of the type /n/ replaces /z/ were significantly more frequent ($F1(1, 29) = 6{\cdot}98$, $p < 0{\cdot}025$; $F2(1, 3) = 8{\cdot}45$, $p < 0{\cdot}10$) than those of the type /z/ replaces /n/. The effect of the prime was significant ($F1(2, 58) = 8{\cdot}86$, $p < 0{\cdot}001$; $F2(2, 6) = 7{\cdot}93$, $p < 0{\cdot}025$) but the interaction was not ($F1$ and $F2 < 1$). Errors were more frequent in the lexico–phonological condition than in the phonological condition ($F1(1, 29) = 9{\cdot}70$, $p < 0{\cdot}005$; $F2(1, 3) = 8{\cdot}22$, $p < 0{\cdot}10$) and than in the control condition ($F1(1, 29) = 10{\cdot}54$, $p < 0{\cdot}005$; $F2(1, 3) = 7{\cdot}88$, $p < 0{\cdot}10$), with no difference being observed between the latter two conditions ($F1$ and $F2 < 1$).

DISCUSSION

One initial result confirms the early availability of /n/ in the liaison errors. A second result reinforces our central hypothesis. In comparison with the control condition, the liaison errors increased when the prime contained the liaison–word2 sequence but not when it contained the liaison alone. For example, when the children heard [zuʀs] in [døzuʀs] they were more likely to produce the error [œ̃zuʀs] whereas hearing /z/ in [døzarbr] did not increase the frequency of the error [œ̃zuʀs]. This higher level of errors in the lexico–phonological condition persisted irrespectively of the type of expected error (the interaction between the priming condition and the target liaison was not significant). Finally, the by-items analysis suggests that the influence of the priming condition was largely independent of the word2s. Despite the small number of items, the F2 associated with this effect was significant and the F2 values associated with the pairwise comparisons were marginally significant.

This pattern of results does not conform to the autosegmental phonology framework. If the children's errors simply resulted from filling an abstract position C with a phonetic content taken from the context, then simply hearing [z] in the prime [døzarbr] should be enough to cause the error [œ̃zuʀs]. On the other hand, this pattern is compatible with the usage-based view that various LCs are attached to the onset of alternative exemplars of the word2. Hearing the exemplar [zuʀs] in the prime activates its representation and increases the likelihood of the error [œ̃zuʀs]. The same process means that hearing the exemplar [nuʀs] provokes the error [dønuʀs]. Although these results seem clear-cut, their interpretation requires two further observations.

The first consideration takes into account results of phonological priming experiments in children. Brooks & MacWhinney (2000) have shown that children aged five to twelve years name pictures faster when the pronounced

https://doi.org/10.1017/S0305000908009124 Published online by Cambridge University Press

word shares the same onset consonant as the heard prime. Observing that the phonological priming does not occur when the auditory stimulus is presented prior to the picture, the authors conclude that this effect concerns the generation of the speech plan: the onset of the heard word primes the activation of a similar onset in the phonological buffer. In the phonological condition of experiment 1, the exemplar /zarbr/, present in the prime [døzarbr], shares the same onset as the intruding exemplar /zuʀs/ that causes the error [œ̃zuʀs]. One could thus expect a priming effect in this condition, which is not the case. In this experiment, the prime is heard before the presentation of the picture. Moreover, we observe the effect of priming on error elicitation and not its ability to shorten the reaction time for the production of correct words. In these conditions, it appears that the results obtained by Brooks & MacWhinney cannot be directly applied to experiment 1. However, further priming experiments should be carried out to specify the organization of the exemplars in the lexicon.

The second observation concerns the well-established fact that the frequency of phonological errors on the initial consonant doubles when the word interacts with another word which shares a vowel with it (Stemberger, 1990). It is therefore possible that each of the children memorized a single, unique variant of each word2 (for example /nuʀs/) and that hearing the other variant in the lexico–phonological condition ([zuʀs]) resulted in the replacement of the initial consonant. For purely phonological reasons, a prime that shares a vowel with the target in the lexico–phonological condition (e.g. *deux ours* as a prime for *un ours*) should lead to the perseveration of the LCs in the output more than a prime that is phonologically distinct (e.g. *deux arbres* 'two trees' in the phonological condition or *deux cochons* 'two pigs' in the control condition). In this case, the liaison errors would not result from the replacement of the exemplar [nuʀs] by the exemplar [zuʀs] but from a well-known phenomenon involving the substitution of initial consonants. However, one argument weakens this alternative interpretation of liaison errors.

A phonological substitution error on the initial element should lead to the initial segment of the word2 being replaced by a wide variety of consonants. For example, *deux ours* could be produced as [døkuʀs] due to the interaction between /zuʀs/ and /kuʀs/ (*course* 'running'). But the LC errors are primarily restricted to replacements by other LCs or by the /l/ which results from elision of the articles *le* and *la*. We reanalyzed the errors collected by Dugua (2006) during a picture naming task involving 200 children aged between two and six years. The word1s and word2s were the same as in experiment 1. In the 1,472 word1–word2 sequences produced by the 200 children, we noted 228 incorrect liaisons, i.e. 15·5% of the productions. Among these errors, 223 (i.e. 97·8%) took the form of the intrusion of the most frequent liaisons or of the /l/ which results from the elision. Only

https://doi.org/10.1017/S0305000908009124 Published online by Cambridge University Press

5 errors (2·2%) involved other consonants and were produced by children aged four years or less (more specifically: /d/, /k/ and /ʁ/). The errors which affected the liaisons were not common phonological errors: disregarding any priming effect, in most cases the liaisons were replaced by other liaisons or by the elision /l/. The incorrect liaisons which children produce are therefore not ordinary phonological substitutions which consist of changing an initial consonant. The most direct way of interpreting the results of experiment 1 is therefore to accept that the errors result from the alternation of word2 exemplars starting with different LCs. One of the goals of experiment 2 is to consolidate this conclusion.

EXPERIMENT 2: IMPERATIVE TASK

The aim of this second experiment was to identify the different exemplars of the word2s and to observe how their availability changes between the ages of two and six years. The task consists of calling to puppets of animals (*Ours, viens ici*! 'Bear, come here!'). Since this task induces the production of word2s that are not preceded by word1s, it tells us about the likelihood that the various exemplars of word2 will be produced when the constraints exerted by word1 are neutralized.

An initial question relates to the production of vowel-initial variants (/urs/ for *ours*, /arbr/ for *arbre* 'tree'). We can ask whether these variants are present at an early stage in the lexicon or whether they stem from a late resegmentation of the consonant-initial exemplars. This second solution was considered by Côté (2005; see also Morel, 1994) who argues that the LCs gradually become detached from the word2 and are then inserted by epenthesis when children have learned to predict the correct LCs as a function of context. The vowel-type variants would be the result of this detachment and the omission errors (*les arbres* pronounced [learbr], without LC) would correspond to an intermediate stage during which detachment occurs even though epenthesis has not yet commenced. Alternatively, we can imagine that the vowel-initial variants of certain word2s are retrieved at an early age in contexts where they are preceded by a word1 which does not induce any liaison (*joli arbre* 'pretty tree'). In this case, the omissions would result from the insertion of the vocalic variant in a schema of type *les*+X.

A second question relates to the production of the consonant-initial variants. In a corpus of 665 errors produced in a natural situation by a small girl aged between 2;1 and 3;7, Chevrot & Fayol (2001) observed that she produced 41 liaisons without a word1 in utterances starting with word2. For example, at 2;10, she referred to an *âne* 'donkey' using a consonant type variant situated at the start of the utterance: [nan]. Our model predicts that such productions should be observed in most children and then disappear when they learn to associate the various exemplars with different contexts.

https://doi.org/10.1017/S0305000908009124 Published online by Cambridge University Press

TABLE 3. *Sample of 200 subjects participating in experiments 2, 3 and 4 (number, age bracket and mean age per age range)*

Age range	Number	Age bracket	Mean
Age range 1	49 children	2;4–3;1	2;9
Age range 2	50 children	3;2–4;1	3;6
Age range 3	52 children	4;2–5;1	4;7
Age range 4	49 children	5;2–6;1	5;7

Subjects

Two hundred children aged between 2;4 and 6;1, distributed over four age ranges (Table 3), took part in this experiment as well as in the following two. Table 3 shows the organization of the sample.

Material and procedure

We used four animal puppets with names which start with a vowel: *âne* 'donkey', *écureuil* 'squirrel', *éléphant* 'elephant', *ours* 'bear'. These names are identical to those used in experiment 1, except that *arbre* 'tree' was replaced by the masculine noun *âne* given that it is unlikely that anyone would want to call to a tree. As distracters, we used three other puppets whose names started with a consonant: *cochon* 'pig', *chien* 'dog', *perroquet* 'parrot'.

At the start of the pass, the child had to choose a puppet from among the distracters and place it in front of him/her. The other six puppets were left on the table, next to one another, facing the one chosen by the child. Adopting the role of the chosen puppet, the child then had to call his or her 'animal friends' by their name in order to summon them to come and join them. The experimenter illustrated the instructions by calling a distracter, for example: *chien, viens ici!* 'Dog, come here!'. Initially, the child called all the puppets and the experimenter moved them forwards as they were called. Once all the puppets had been called, the same task was restarted in order for the child to produce eight vowel-initial nouns, with each of the puppets corresponding to the donkey (*âne*), squirrel (*écureuil*), elephant (*éléphant*) and bear (*ours*) being called twice. All of the children decided on the animals that they wanted to call, with the order of production varying between subjects.

RESULTS

In addition to the expected consonant-initial and vowel-initial productions, we observed productions with a definite or indefinite determiner – *l'ours*

https://doi.org/10.1017/S0305000908009124 Published online by Cambridge University Press

TABLE 4. *Experiment 2 – Imperatives task* (*mean percentages and standard deviations*)

Age ranges	Vowel-type variants (e.g. *ours!*)	LC-initial variants (e.g. *nours!*)	Productions with determiner (e.g. *un ours! l'ours!*)
Range 1 2;4–3;1	40·6% (32·9)	35·4% (32·7)	24·0% (33·8)
Range 2 3;2–4;1	47·5% (37·1)	19·0% (27·2)	33·4% (33·6)
Range 3 4;2–5;1	83·3% (25·3)	8·0% (15·3)	8·7% (23·4)
Range 4 5;2–6;1	81·5% (26·0)	2·8% (10·8)	15·8% (24·3)
Total 2;4–6;1	64·3% (36·0)	15·5% (25·6)	20·1% (30·2)

'the bear' or *un ours* 'a bear'. We considered that such cases did not constitute utterances starting with a word2. As you can see in Table 4, these responses are presented in the four age ranges and represent a total of 20·1% of all the productions.[3] When all age ranges are taken together, the adult-like vowel-initial variants (*ours!* [urs] 'bear!') represent 64·3% of all the productions. The variants starting with /n/, /z/ or /t/ (LC-initial productions) represented 15·5% of the produced utterances.

Among the LC-initial variants, productions starting with /n/ were more frequent (93%) than those starting with /z/ or /t/ (7%). This difference was observed in all the age groups. Even the distribution in which the levels of /z/ and /t/ liaisons were highest (14% at four years of age) exhibited a great preponderance of /n/ liaisons (86%). We were therefore again able to note the high level of /n/ liaisons in the errors. Furthermore, the presence of a consonant other than /l/, /z/, /n/ or /t/ at the start of the noun was marginal. Of the 1,347 nouns produced with apostrophe by the 200 children, we noted 373 occurrences starting with a consonant (/l/, /z/, /n/, /t/, /j/, /s/, /d/, /ʀ/). Of these 373 cases, 363 (97·3%) involved the presence of one of the LCs or the presence of /l/. Only ten errors (2·7%) involved other consonants (/t/, /j/, /s/, /d/, /ʀ/) and were observed in 8 children aged less than four years. As in experiment 1, the errors recorded in experiment 2 cannot be considered as ordinary phonological additions since they primarily involved the three LC's and the /l/ which results from article elisions.

[3] When we speak of 'all productions', we mean the productions minus non-responses (the child said nothing), errors in which the child cut off the initial vowel and those in which the child named the wrong word (*cheval* 'horse' instead of *âne* 'donkey').

https://doi.org/10.1017/S0305000908009124 Published online by Cambridge University Press

Due to the heterogeneous nature of the variances, non-parametric tests were used in place of an ANOVA[4] to analyze the individual scores for vowel-initial productions and LC-initial productions. Vowel-initial productions did not change between age ranges 1 and 2 (Mann–Whitney $U = 948{\cdot}5$, $p = 0{\cdot}4059$), increased significantly between the next two ($U = 580$, $p < 0{\cdot}0001$) and remained stable between age ranges three and four ($U = 1242{\cdot}5$, $p = 0{\cdot}6694$). LC-initial productions fell significantly between all consecutive age ranges (between ranges 1 and 2: $U = 761{\cdot}5$, $p = 0{\cdot}0148$; between ranges 2 and 3: $U = 1015{\cdot}5$, $p = 0{\cdot}0350$; between ranges 3 and 4: $U = 1080$, $p = 0{\cdot}0232$).

If we look at the differences within each age range, we see that the percentages corresponding to the vowel-type variants and those corresponding to the LC-initial variants were not significantly different in age range 1 (Wilcoxon, $z = -0{\cdot}524$, $p = 0{\cdot}6002$). In age ranges 2, 3 and 4, the vowel-type variants were systematically more frequent than the consonant-type variants (range 2: Wilcoxon, $z = -3{\cdot}126$, $p = 0{\cdot}0018$; range 3: $z = -5{\cdot}968$, $p < 0{\cdot}0001$; range 4; $z = -6{\cdot}013$, $p < 0{\cdot}0001$).

Finally, an observation of the individual scores in each age range shows that 43 out of the 49 children in age range 1 produced at least 1 LC-initial variant. This proportion was maintained in range 2 (40/50) and then fell in the latter two ranges (21/52 and 24/49 respectively). Second, 30 children out of 49 produced at least one vowel-type variant in range 1, and this proportion increased in ranges 2, 3 and 4: 38/50, 50/52, 49/49. Third, at all ages, a large proportion of the children produced both types of variants: 30/49, 29/50, 19/52 and 23/49 in ranges 1, 2, 3 and 4 respectively.

CONCLUSION

The production of LC-initial variants at the start of an utterance is not specific to the case observed by Chevrot & Fayol (2001). These variants are produced by the majority of children up to the age of four years, with their frequency then diminishing continuously even though half of the subjects continue to produce a small number of them at the age of five to six years. Throughout this period of development, variants starting with /n/ are more readily available than those starting with /z/ or /t/. These results reinforce those of experiment 1: in the early lexicon, the LC are located at the start of the word2s. Since the errors persist when the word2 is produced without a word1, the possibility of a post-lexical resyllabification of an LC attached to the word1 is totally excluded.

[4] More than the non-normality of the distributions, the homogeneity of the variances is a crucial condition for the validity of the ANOVA (Judd, McLelland & Culhane, 1995).

https://doi.org/10.1017/S0305000908009124 Published online by Cambridge University Press

Vowel-initial variants are observed in more than half of children aged two to three years, the age in which their frequency is not significantly different from the LC-initial variants. Their occurrence increases rapidly between three and four years and they become more frequent than the consonant-type variants with all children producing them at the age of five to six years. The early presence of vowel-type variants is difficult to reconcile with the idea that they result from a late resegmentation of LC-initial variants when children learn the correct liaisons (Côté, 2005; Morel, 1994). At two to three years, the vowel-type variants are already available, whereas only 36% of obligatory liaisons between determiner and noun are correct in a naming task, as against 80% at four to five years (Dugua, 2002: 62). We therefore support the idea that the vowel-type variants are retrieved from the input and that their insertion in schemas of type *deux*+X results in omission errors.

While the two types of variants are present at two to three years, they subsequently develop in different ways. The use of the LC-initial variants at the start of an utterance declines progressively whereas the use of the vowel-type variants increases rapidly at about three to four years. This suggests that a generalization process is involved in the ability to associate exemplars with contexts. The aim of the following experiments was to describe this process of generalization.

EXPERIMENT 3: PSEUDOWORD SEGMENTATION

In the developmental scenario proposed here, production is underpinned by two construction systems. The earlier system consists of schemas of type *un*+X or *deux*+X which contain no information about the nature of the liaison triggered by the word1. Because of this absence of information, children randomly insert one of the word2 exemplars (/nurs/, /zurs/, /urs/ for *ours* 'bear') in position X. Depending on the case in question, they produce substitutions (*deux*+/nurs/), omissions (*deux*+/urs/) or correct sequences (*deux*+/zurs/). The second system appears later and protects speakers against omission and substitution errors. It consists of schemas which generalize the relation between a word1 and the word2 exemplars which start with an LC triggered by this word1. For example, the schema *un*+/nX/ establishes a relation between *un* and the exemplars which start with /n/ while the schema *deux*+/zX/ associates *deux* with exemplars which start with /z/.

This second system has to be productive. In effect, it is unlikely that a child will have encountered, segmented and memorized all the exemplars of all the word2s in French (for example, the /n/, /z/ and /t/ exemplars of each of the French nouns which start with a vowel in adults). As a result, a child who wants to produce the sequence *un ami* 'a friend' but does not possess

https://doi.org/10.1017/S0305000908009124 Published online by Cambridge University Press

the exemplar /nami/ which is required by the schema *un*+/nX/ will have to create it by analogy with the complete paradigms that he or she has memorized. This possibility is attested to by anecdotal accounts of errors of the type [œ̃nɛbr] (instead of *un zèbre* 'a zebra') in which the child replaces the fixed initial /z/ of *zèbre* by the consonant /n/, which is precisely the LC induced by *un* (Dugua, 2006; Wauquier-Gravelines, 2003). We consider these errors to represent an overgeneralization of an analogue process which enables children to generate all the word2 variants on the basis of the memorized exemplars.

The aim of experiment 3 was to describe the developmental evolution of the two construction systems. The task consisted of asking the children to listen to pseudo-nouns preceded by a determiner which induces an LC and then asking them to produce the same pseudo-noun in association with a determiner which requires a different LC. On the auditory presentation of the pseudo-noun, the preceding determiner was either the singular *un* 'one' which requires an /n/ liaison, or the plural *deux* 'two' which requires a /z/ liaison. In the singular case, the consonant heard by the children between the determiner and the pseudonoun was /n/: for example, the child heard *un-n-ivak* [œ̃nivak]. In the plural case, this consonant was /z/: for example, the child heard *deux-z-ikat* [døzikat]. The child then had to extract the pseudonoun heard in *un-n-ivak* and insert it in the context *deux*__ which requires the LC /z/; alternatively, the child had to extract the pseudonoun heard in *deux-z-ikat* and then produce it in the context *un*__ which requires the LC /n/.

During the auditory presentation, the children did not possess adequate cues permitting them to decide whether the /z/ or the /n/ heard in *deux-z-ikat* or *un-n-ivak* were LCs or fixed initial consonants as at the start of *zèbre* 'zebra' or *nombril* 'navel'. In effect, if adults are asked to decide whether the acoustic sequence [œ̃nœf] corresponds to *un œuf* 'an egg' with a liaison in /n/ or to *un neuf* 'a nine' with an initial /n/, their responses do not differ from chance (Yersin-Besson & Grosjean, 1996). Even if acoustic cues are involved in the lexical processing of the words which follow the LC (Spinelli, McQueen & Cutler, 2003), the lexical information is necessary in order to resolve the boundary ambiguity (Spinelli & Meunier, 2005). In the case of pseudo-words or new words, this lexical information is lacking. An input such as *un-n-ivak* [œ̃nivak] does not therefore make it possible to specify whether the noun which follows the determiner *un* is *ivak* [ivak] or *nivak* [nivak]. It must therefore be accepted that children are able to retrieve a vowel-initial variant or even a consonant-initial variant on the basis of this type of input. In the predictions made concerning experiment 3, we take account of this uncertainty by listing the sequences which result from inserting one or other of the two variants in the first or second construction system.

https://doi.org/10.1017/S0305000908009124 Published online by Cambridge University Press

The use of the first system results in two types of sequences. Depending on whether the children retrieve *nivak* or *ivak* from the heard sequence *un-n-ivak*, they produce either[dønivak] or [døivak] by inserting the retrieved variant in the schema *deux*+X. Thus the first construction system necessarily results in the production of sequences which either retain the heard /n/ or /z/ or lead to their omission. The mobilization of the second system results in a third type of production. In effect, whatever the variant retrieved from the heard sequence *un-n-ivak* ([nivak] or [ivak]), children create a sequence which is compatible with the schema *deux*+/zX/ which requires a variant starting with /z/. The use of the second system thus results in the replacement of the heard /n/ by a /z/ (alternating responses). We therefore predict that the omissions or retentions of the heard consonant, which are typical of the first schema, will account for the majority of responses at an early age and will then tend to disappear, while alternating responses, which are typical of the second system, should gradually come to represent the majority of responses.

Subjects

The subjects who took part in these task were the same 200 children aged 2;4 to 6;1 who participated in experiment 2.

Material and procedure

We used four pseudo-words to designate imaginary animals drawn in pictures. These pseudo-words took the form of masculine bisyllabic words: [ivak], [ytRɛl], [ikat], [ysa]. None of them had any nominal phonological neighbour in French either in the vowel-initial form ([ivak]) or in a consonant-initial form resulting from a liaison or an elision ([nivak], [zivak], [tivak], [livak]). The auditory presentation of the pseudowords and their production after the determiners *un* 'one' and *deux* 'two' took the following form. Initially, the experimenter presented the picture of the animal and said, for example: *Sur cette image, il y a un-n-ivak* [œ̃nivak] 'In this picture, there is [œ̃nivak]'. Next, he showed the child a picture depicting the same animal drawn twice and the child had to say what he or she could see. In contrast, if the auditory presentation of the pseudo-word followed the determiner *deux*, the child had to produce it after the determiner *un*.

We divide the pseudo-words into two groups: the first consisted of [ivak] and [ytʀɛl], and the second of [ikat] and [ysa]. Half of the children heard the pseudo-words of the first group preceded by *un* and the pseudo-words of the second group preceded by *deux*, with this presentation being inverted

https://doi.org/10.1017/S0305000908009124 Published online by Cambridge University Press

TABLE 5. *Experiment 3 – Pseudo-word segmentation task (mean scores and standard deviations)*

Age ranges	Retention of /n/ or /z/ Un-**n**-ikat→[dø**n**ikat]	Alternation n/z Un-**n**-ikat→[dø**z**ikat]	Omission of /n/ or /z/ Un-**n**-ikat→[døikat]
Range 1 2;4–3;1	2·64 (1·32)	0·60 (0·86)	0·45 (0·83)
Range 2 3;2–4;1	2·66 (1·24)	0·68 (1·06)	0·40 (0·85)
Range 3 4;2–5;1	2·03 (1·39)	1·69 (1·38)	0·25 (0·59)
Range 4 5;2–6;1	1·40 (1·16)	2·28 (1·37)	0·24 (0·59)
Total 2;4–6;1	2·13 (1·37)	1·39 (1·39)	0·32 (0·71)

for the other half of the children. Using this procedure, no child heard the same pseudo-word preceded by *un* and preceded by *deux*. The order of presentation of the four pseudo-words was random but alternated between the two groups, i.e. between the two numbers.

RESULTS

An initial result relates to the non-responses. In effect, out of the four expected productions, 16 children in age range 1 and 3 children in age range 2 produced 2 or more non-responses. Because of the impact on the validity of the scores, these subjects have been removed from the following analyses.

Table 5 indicates that the level of retention of the heard consonant fell between age ranges 2 and 3 (Mann-Whitney, $z = -2{\cdot}165$, $p = 0{\cdot}03$) and between ranges 3 and 4 (Mann-Whitney, $z = -2{\cdot}309$, $p = 0{\cdot}02$). In contrast, the alternating responses increased between ranges 2 and 3 (Mann-Whitney, $z = -3{\cdot}612$, $p = 0{\cdot}0003$) and again between ranges 3 and 4 (Mann-Whitney, $z = -2{\cdot}085$, $p = 0{\cdot}037$). Due to this crossed developmental pattern, the mean retention level was greater than the mean alternation level at age range 1 (Wilcoxon, $z = -4{\cdot}094$, $p < 0{\cdot}0001$) and at age range 2 (Wilcoxon, $z = -4{\cdot}513$, $p < 0{\cdot}0001$). At age range 3, the difference was no longer significant ($p = 0{\cdot}346$). At range 4, alternating responses were more prevalent (Wilcoxon, $z = -2{\cdot}527$, $p = 0{\cdot}0115$).

Omission-type responses were less frequent than the other two types. With all ages taken together, the mean value for such responses was 0·32 whereas retentions reached 2·13 and alternations 1·39. Finally, even if the curve for omissions tended to fall, the differences between consecutive ages

https://doi.org/10.1017/S0305000908009124 Published online by Cambridge University Press

or between the two extreme ages did not reach significance (Mann-Whitney, $p > 0{\cdot}41$ whatever the comparison).

DISCUSSION

Retentions of the heard consonant are the manifestation of the first construction system consisting of schemas of type *un*+X or *deux*+X which contain no information about liaison. The alternating responses result from the mobilization of the second system which consists of schemas of type *un*+/nX/ or *deux*+/zX/ which link a word1 to a class of variants starting with a specific LC. The development changes in the number of retentions and alternations supported the predictions: the first system is available at an early age and the second system becomes operational later.

The omissions result from the convergence of two operations. First, children extract a vowel-initial variant after hearing the pseudo-noun: for example /ivak/ on the basis of *un-n-ivak* ([œ̃nivak]) or /ikat/ on the basis of *deux-z-ikat* ([døzikat]). Second, they insert the extracted variant into one of the constructions from the first system: /ivak/ is inserted into *deux*+/X/ and /ikat/ into *un*+X. Contrary to the predictions, the omissions did not decline with age even though they are an index of the availability of the constructions from the first system. This result suggests that these constructions remain just as available at five to six years as they are at two to three years when children are required to insert vowel-initial variants.

This possibility of inserting vowel-initial variants in the constructions coming from the first system can be seen not only in pseudo-words but also in known words. In effect, in a task requiring the naming of genuine French words after the determiners *un* and *deux*, Dugua, Chevrot & Côté (2003) observed that the liaison omission rate varied from 20% at two to three years to 14% at five to six years without there being any significant difference in the means for the age ranges. This result will be discussed in the general discussion.

EXPERIMENT 4: NAMING WORDS OF TYPE *ZÈBRE*

The productivity of the second construction system is seen in errors which consist of replacing a consonant at the start of a word with an LC which is compatible with the preceding word: production of [œ̃nɛbr] for the sequence *un zèbre* 'a zebra'. Since these are generalizations, their change during development should be similar to that of the errors observed, for example, during the acquisition of verbal morphology: an increase followed by a decrease. To date, these errors have been recorded only sporadically during case studies (Dugua, 2006; Wauquier-Gravelines, 2003). The aim of experiment 4 was to gather systematic data to describe their developmental pattern and assess the proportion of children in whom they occur.

https://doi.org/10.1017/S0305000908009124 Published online by Cambridge University Press

Material, procedure, subjects

We designed a naming task including four masculine nouns: *nombril, nuage, zèbre, lavabo* 'navel', 'cloud', 'zebra', 'washbasin'. The specific characteristic of these nouns is that they have an initial consonant which, in another context, could be a liaison (/n/ or /z/) or the consonant /l/ resulting from the elision of the articles *le* and *la* 'the'. Each of these nouns is involved in the errors which Dugua (2006) noted in the utterances produced by a young girl between the ages of 2;1 and 6;4. Furthermore, each of these nouns corresponds to an object or animal which can be easily represented in a picture.

One or two exemplars of each of the four nouns were depicted in a series of pictures which the children had to name. Thus, the picture naming task elicited the production of the four target nouns after the determiner *un*, which induces an /n/ liaison, and after the determiner *deux*, which triggers a /z/ liaison. Even though the order of presentation was random, we nevertheless made sure that the same noun was not produced twice in succession. The subjects were the same as in experiments 2 and 3.

RESULTS

Four types of errors affecting the initial consonants of the words *nombril, nuage, zèbre, lavabo* were observed. First, in the word1-compatible replacements, the children replaced the initial consonant of the noun with the LC triggered by the preceding determiner. For example, they said *un nèbre* ([œ̃nɛbʀ], 'a/one zebra') instead of *un zèbre* ([œ̃zɛbʀ]), with the replacement /n/ being precisely the LC that the determiner *un* triggers; or they said *deux zombrils* ([døzɔ̃bril], 'two navels') instead of *deux nombrils* ([dønɔ̃bril]), where the /z/ is the LC that the determiner *deux* triggers, etc. These productions constituted the generalization errors that we expected. Second, in the word1-incompatible replacements, the children replaced the /z/ of *zèbre* by an /n/ after the determiner *deux* ([dønɛbʀ]) even though /n/ is not the LC that this determiner triggers. Third, in cases where the initial consonant was omitted, the children spoke the sequence *un zèbre* without the initial /z/ of word2: [œ̃ɛbr]. Fourth, we also observed cases in which the initial consonant of the target words was replaced by a /l/ and not by a consonant which can act as a liaison (i.e. /z/ or /n/). In such cases, the children produced *un lèbre* ([œ̃lɛbr]) instead of *un zèbre*.

The mean individual percentage values[5] and the total number of errors for each type and each age range are presented in Table 6. It can be seen

[5] These percentages are measured against the number of actual responses. Cases in which the children produced nothing or produced an unexpected determiner or noun (*un cheval* 'a horse' instead of *un zèbre* 'a zebra') were excluded from the denominator.

https://doi.org/10.1017/S0305000908009124 Published online by Cambridge University Press

TABLE 6. *Experiment 4 – Naming words of type* zèbre (*mean percentages of errors and total number of errors*)

Age ranges	Replacement of the initial C by an LC compatible with the word1 *un nèbre* [œnɛbr]	Replacement of the initial C by an LC NOT compatible with the word1 *deux nèbres* [dønɛbr]	Omission of the initial consonant *un èbre* [œɛbr]	Replacement of the initial C by /l/ *un lèbre* [œlɛbr]
Range 1	5%	3%	1%	3%
2;4–3;1	5	2	2	4
Range 2	2%	0·3%	3%	2%
3;2–4;1	6	1	6	4
Range 3	11%	0·5%	4%	0%
4;2–5;1	40	2	9	0
Range 4	5%	0·4%	4%	0·7%
5;2–6;1	15	1	9	2
Total	6%	1%	3%	1%
2;4–6;1	66	6	26	10

that only the replacements compatible with the word1 had a significant impact. Without differentiating between the age ranges, their percentage of occurrence was greater than that of the other types (whatever the pairwise comparison using the Wilcoxon test, $z \leqslant -2{\cdot}574$, $p \leqslant 0{\cdot}01$) but the other types do not differentiate between pairs (Wilcoxon, $z \geqslant -1{\cdot}378$, $p \geqslant 0{\cdot}1683$). Furthermore, the replacements compatible with word1 changed with age (Kruskal–Wallis, $H = 16{\cdot}114$, $p < 0{\cdot}0011$) whereas this was not the case for any of the other three types of error (irrespective of type, Kruskal–Wallis, $H \leqslant 0{\cdot}731$, $p \geqslant 0{\cdot}8659$).

This change in replacements compatible with the word1 can be seen more clearly when we compare them with the non-compatible replacements. Neither of these two error types changed between age ranges 1 and 2: the apparent decrease was not significant either for the compatible errors (Mann–Whitney, $p = 0{\cdot}6035$) or for the incompatible errors ($p = 0{\cdot}3774$). From the third age range onwards, the two curves diverge. The non-compatible errors stagnate at a very low level and the differences between the successive age ranges are not significant ($p \geqslant 0{\cdot}5675$). The compatible errors increase significantly between age ranges 2 and 3 ($U = 757{\cdot}000$, $p < 0{\cdot}0001$) and then decrease significantly between age ranges 3 and 4 ($U = 907{\cdot}000$, $p = 0{\cdot}0024$).

In order to assess the proportion of children producing word1-compatible replacements in each age range, we restricted our analysis to those subjects who produced at least two of the eight expected determiner + noun sequences. In effect, the high level of non-responses and lexical substitutions (*un âne* 'a donkey' instead of *un zèbre* 'a zebra') observed in the youngest subjects reduced the probability of this type of error being

https://doi.org/10.1017/S0305000908009124 Published online by Cambridge University Press

TABLE 7. *Experiment* 4 – *Naming words of type* zèbre*: proportion of children who made word*1*-compatible initial consonant replacement errors*

Age ranges	Total number of children	Number of children producing at least two determiner–noun sequences	Proportion of children producing at least one word1-compatible replacement
Range 1 2;4–3;1	49	27	18·5% (5/27)
Range 2 3;2–4;1	50	45	11·1% (5/45)
Range 3 4;2–5;1	52	52	51·9% (27/52)
Range 4 5;2–6;1	49	49	22·4% (11/49)

produced. For each age range, Table 7 indicates the total number of children, the number of children who produced at least two expected sequences and the proportion of such children who produced at least one word1-compatible replacement.

It can be seen that the proportion of children producing the expected errors increased sharply at around age four to five years and then fell again at five to six years. At four to five years, all the children produced at least two target determiner–noun sequences and more than half replaced the initial consonant of the words *nombril, nuage*, *zèbre, lavabo* with a word1-compatible LC at least once.

DISCUSSION

Errors consisting of the replacement of initial consonants with an /n/ or /z/ compatible with the word1 are not specific to the children observed in the case studies conducted by Dugua (2006) and Wauquier-Gravelines (2003). These errors were also observed between determiners and nouns in more than half of the four- to five-year-old children who took part in experiment 4. Their individual frequencies and the number of children who produced them culminated at about four to five years, the age at which the level of correct liaisons between determiners and nouns rose to above 80% (Dugua, 2006). This co-occurrence of the generalization of correct liaisons and the peak error rate strongly suggests that we are witnessing a process of regularization.

As mentioned above, we consider this pattern of errors to be an indication that the item-based constructions of the type *deux* + /zX/ linking a word1 to a set of /z/-initial word2 exemplars are becoming productive. Due to this productivity, children modify the known form *nuage* into *zuage*, in order to fit the schema based on the word *deux*. However, another possibly more

https://doi.org/10.1017/S0305000908009124 Published online by Cambridge University Press

direct interpretation exists (Wauquier-Gravelines, 2003). This consists of accepting that these errors occur when children attach the LC to the word1s. The lexical representations of the determiners *un* and *deux* are then /œn/ and /døz/, terminating with /n/ and /z/. When children are confronted with the segmentation of the word *nuage* in the sequence /œnya3/ (*un nuage* 'a cloud'), they start with the known lexical form /œn/ and position the word boundary after the /n/. They therefore arrive at the incorrect segmentation /ya3/ for the word *nuage*. If they re-use this form /ya3/ after the determiner /døz/ equipped with the final LC /z/, they produce what we have termed a word1-compatible replacement error: /døzya3/. If this hypothesis is correct, this type of error would not reflect the productivity of schemas of the type *deux*+/zX/ but instead the emerging ability to encode the LC at the end of the word1s and to segment vowel-initial word2s. In this case, the alternating segmentations of the determiner+pseudonoun sequences in experiment 3 would have a different significance. In effect, the production of [døzivak] after hearing *un-n-ivak* ([œnivak]) would primarily indicate the ability to retrieve the form [ivak] and combine it with a lexical representation of the determiner *deux* equipped with a final /z/. It is therefore essential to choose between these two possible interpretations of the errors. To this end, we conducted an additional experiment based on a judgment of acceptability task.

EXPERIMENT 5: JUDGMENTS OF THE ACCEPTABILITY OF WORDS OF TYPE *ZÈBRE*

If an incorrect segmentation leads children to memorize the form /ya3/, then they should accept the sequence [ya3] as a variant of the word *nuage*. In contrast, if the error /døzya3/ is due to a modification of the memorized form /nya3/ in order to fit the schema *deux*+/zX/, then children should accept [nya3] as a variant of the word *nuage*.

Subjects

This experiment involved 123 children aged between 2;9 and 6;3, an age range which includes the error peak observed in experiment 4. None of them had taken part in the preceding experiments.

Material and procedure

These subjects performed two tasks. First of all, they had to name, in random order, sixteen pictures which required them to produce twice, each of the combinations of the determiners *un* and *deux* with the nouns *nombril, nuage*, *zèbre, lavabo* 'navel', 'cloud', 'zebra', 'washbasin'. They then took

https://doi.org/10.1017/S0305000908009124 Published online by Cambridge University Press

part in a judgment of acceptability task. They heard each of the words from the production task in isolation and in two forms, i.e. in their correct form ([nyaʒ]) and without the initial consonant ([yaʒ]). The two forms were spoken by two puppets which were controlled by the experimenter, who pointed to the picture of a cloud and, speaking normally, made one puppet say [nyaʒ] and the other say [yaʒ]. The child had to point to the puppet which he or she thought had spoken the word correctly. Each pair of words corresponding to the four words was presented twice in each of the possible orders ([nyaʒ]–[yaʒ] and [yaʒ]–[nyaʒ]), with each child making sixteen judgments, i.e. four for each word. Our analysis consisted of observing how the children who made errors during the production of a specific noun judged this same noun. This amounts to asking the following question: do children who make the error /døzyaʒ/ know that the word *nuage* starts with /n/ or do they think that the word is /yaʒ/?

RESULTS

During the production task, sixteen children made errors which consisted of replacing the initial /l/ of *lavabo* with an /n/ or /z/ compatible with the preceding determiner. During the judgment task, these children chose the form [lavabo] 2·75 times on average and the form [avabo] only 0·93 times. These two means are significantly different from the random value 2 (bilateral, univariate *t*-tests, $p < 0{\cdot}045$), so neither the choice of [lavabo] nor that of [avabo] can result from chance. The judgments were identical in the nine children who made production errors on *nombril*. The mean judgment value corresponding to the form [nɔ̃bril] was 3·22, while that corresponding to [ɔ̃bril] was 0·79 (bilateral univariate *t*-tests compared to random value 2, $p = 0{\cdot}01$). The same was true of the eight children who made errors on *zèbre*: mean score for the choice of [zɛbr]: 3·5; mean score for choice of [ɛbr]: 0·5 (bilateral univariate *t*-tests, $p = 0{\cdot}0008$). As far as the last word is concerned, the form [nyaz] ($m = 2{\cdot}8$) was chosen considerably more often than the form [yaʒ] ($m = 0{\cdot}4$), but only this second mean was significantly different from the random value 2 (*t*-test for [nyaʒ]: $p = 0{\cdot}37$, *t*-test for [yaʒ]: $p < 0{\cdot}02$). It should be emphasized that the number of subjects was small since only five children made production errors on *nuage*. However, three of them had a maximum score of four judgments in favour of [nyaʒ]. They therefore produced the error [døzyaʒ] even though they were certain that *nuage* starts with the consonant /n/. Since the same was observed for the other three words, we conclude that the word1-compatible replacements were not due to incorrect segmentations such as /yaʒ/. Instead, these errors result from the productivity of schemas of type *deux*+/zX/ whose implementation during production results in the on-line transformation of the memorized form /nyaʒ/ into [zyaʒ].

https://doi.org/10.1017/S0305000908009124 Published online by Cambridge University Press

GENERAL DISCUSSION

In the linguistics field, liaisons are often used as indicators of interactions between the various levels of language organization: phonology, lexicon, morphology, syntax and sociolinguistic aspects (Chevrot, Fayol & Laks, 2005). The current study pursues the same goal but in the developmental field. It presents and documents a model which weaves together several issues: progress in the production of liaisons, the segmentation of new words when the lexical boundaries are unclear, the long-term stabilization of the word form in the lexicon, and the formation of item-based constructions.

The analysis of the child errors (discussion of experiment 1 and results of experiment 2) strongly suggests the specificity of liaisons. Indeed, at least 97% of the substitutions affecting liaisons take the form of another liaison or the elision /l/. Cases in which a liaison is replaced by a segment other than /n/, /z/, /t/, /l/ occur only sporadically in the youngest children. Moreover, at the age of two to three years, less than 50% of the liaisons between determiner and noun are produced correctly in a denomination task (Chevrot, Nardy, Barbu & Fayol, 2007; Dugua, 2006) whereas 70% of French consonants are correctly produced at the age of two years (Vinter, 2001). Thus, it does not seem possible to consider the LCs as ordinary consonants subject to ordinary phonological substitutions.[6] The high level of liaison errors once again emphasizes their specific nature which we take into account by supporting the hypothesis of an early system of alternating exemplars.

In experiment 1, the results of a task which induced errors by means of intermodal priming in fact suggest that children aged three to four years possess multiple lexical exemplars of the words which follow liaisons in speech, with each of these exemplars starting with a specific LC (for example /nami/ and /zami/ for the word *ami* 'friend'). This result strengthens the central idea of stage 1 of the developmental scenario: when children segment word1–LC–word2 sequences (*les amis* [lezami] 'the friends'), they favour initial CV syllables. They therefore attach the liaison to the start of the lexical representation of word2 (/zami/). Since they encounter each word2 preceded by different LCs, the maintenance of the CV segmentation in each context results in the memorization of various alternating exemplars of the same word2 (/zami/, /nami/, etc.).

[6] Nor is it possible to consider that the liaison errors result from the syllabic enchaînement which enables all final consonants to form a syllable with the following vowel (*il arrive à pied*: [i.la.ri.va.pje] 'he comes by foot'). Indeed, the case of enchaînement is different in the sense that the linking consonant occurring at the end of the word1 is realized in all the contexts (before consonant, vowel and pause). The link between this consonant and the word1 is therefore stronger.

https://doi.org/10.1017/S0305000908009124 Published online by Cambridge University Press

This conclusion is supported by a second piece of evidence. Chevrot *et al.* (2007) observed 898 liaison errors produced by a small girl aged between 2;1 and 6;4. A word-by-word reanalysis of the data shows that 122 word2s were involved in these errors, i.e. an average of 7·3 occurrences for each word2. Of the 28 word2s with an above-average level of occurrence (7·3), 21 were preceded by at least two different consonants and 9 by at least three different consonants. These results consolidate the observation that one and the same child produced the word2 *orage* 'thunderstorm' preceded by /l/, /n/ and /z/ within a period of just a few minutes (Chevrot & Fayol, 2001).

In experiment 2, an imperative task showed that the majority of children aged four years or less produce at least one consonant-initial variant at the start of an utterance ([nan] instead of [an] to say *Ane!* 'Donkey!') and that the production of such variants persists through to the age of five to six years. Word2 exemplars starting with an LC are therefore certainly available in children's lexicons. From the age of two to three years, the vowel-initial variants ([an] for *Ane!*) are also available since they are as frequent as the consonant-initial variants. They become predominant at about three to four years.

The aim of experiment 3 was to test the availability of two item-based construction systems. The first system (stage 1) consists of general schemas which include no information about the liaison which follows a word1 (*un*+X or *deux*+X). The constructions formed in the second system (stage 2) link each word1 to a class of word2 exemplars starting with a specific LC (*un*+/nX/ or *deux*+/zX/). The task consisted of extracting a pseudo-noun from a heard sequence and then producing it with a determiner of opposite number. The results suggest that as of the age of five to six years, constructions which contain information about liaisons become more available than those which do not contain this information. We also noticed a slight but persistent tendency to extract a vowel-initial variant and include it in the constructions issuing from the first system (for example, extracting [ikat] on hearing *un-n-ikat*, inserting this form in the schema *deux*+X and producing the sequence [døikat] without an LC). Since the late persistence of liaison omission errors was also observed between determiners and real nouns (Dugua, 2002; Dugua *et al.*, 2003), we must accept that the X slot in the constructions from the first system can be occupied either by a vowel-initial or by a consonant-initial exemplar. This variability raises the general question of the relations that are established between exemplars and constructions.

Conventionally, we think of constructions as generalizations performed on the basis of relations established between memorized sequences, without this generalization necessarily deleting the concrete expressions which lie at its root (Kemmer & Barlow, 2000; Tomasello, 2003: 106). Within this conception, the schema *des*+/X/ emerges at an early age from the

https://doi.org/10.1017/S0305000908009124 Published online by Cambridge University Press

establishment of connections between sequences such as [defij], [dezami], [dezɛbr] (*des filles* 'girls', *des amis* 'friends', *des zèbres* 'zebras'). At the same time, the schema *un*+X results from hearing sequences containing the determiner *un*. In young children, the X slot of these schemas may be occupied by exemplars starting with any vowel or any consonant. Later, when children have stored a sufficient number of sequences which exhibit the correct relation between *des* and /z/ ([dezarbr], [dezami], [dezɛbr], *des arbres* 'trees', *des amis* 'friends', *des zèbres* 'zebras') or between *un* and /n/ ([œ̃narbr], [œ̃nami], [œ̃nɔ̃bril], *un arbre* 'a tree', *un ami* 'a friend', *un nombril* 'a navel'), they develop a second system of schemas which generalize these relations (*des*+/zX/ and *un*+/nX/). When this second system is available, the segmented exemplars starting with /n/ are inserted in a schema *un*+/nX/ and the segmented exemplars starting with /z/ are inserted in the schema *des*+/zX/. As a result, the liaison substitution errors tend to disappear. However, the schemas in the first system remain active. By default, they accept the exemplars which start with any consonant other than /z/ or /n/ as well as those starting with a vowel. In this latter case, the result is a word1–word2 sequence without a liaison. We would therefore expect liaison omissions to persist while substitutions disappear.

The aim of experiment 4 was to identify the frequency and developmental evolution of errors affecting nouns which, in adult speech, start with a fixed /n/ or /z/ (*des nombrils* 'navels' pronounced *des zombrils*). These errors are indicators of the productivity of the stage 2 constructions which include liaison-related information. In effect, children create exemplars starting with /n/ in order to satisfy the requirements of schemas of the type *un*+/nX/ and exemplars of the type /z/ to satisfy the requirements of schemas of the type *des*+/zX/. The results of a naming task which required the production of the determiners *un* and *deux* followed by the nouns *nombril, nuage*, *zèbre, lavabo* ('navel', 'cloud', 'zebra', 'washbasin') showed that the pattern of errors is identical to that of the regularizations observed in morphological development: their individual frequency and the proportion of individuals who make such errors both peak at around four to five years before falling again.

Finally, the judgment of acceptability task in experiment 5 weakened the alternative hypothesis which holds that these errors are due to incorrect segmentation of the word2 (*nombril* would be memorized /ɔ̃bril/ and then used after the determiner *des* accompanied by a final LC /z/). In effect, the children who make these errors know perfectly well that the word2s in question start with an /n/ or a /z/ and not with a vowel. We therefore maintain that the error [dezɔ̃bril] is produced when the schema *des*+/zX/ modifies the memorized form /nɔ̃bril/.

The developmental scenario underpinned by this set of results raises a number of theoretical questions which we discuss below.

https://doi.org/10.1017/S0305000908009124 Published online by Cambridge University Press

In our opinion, the segmentation of lexical items and the formation of morphosyntactic schemas are two aspects of the same developmental process. It is by establishing relations between memorized chunks that children simultaneously succeed in constructing an independent representation of the noun and are able to generalize the determiner-based schemas. When the stock of memorized chunks increases and changes, the possibilities for generalization develop. Consequently, the schemas which specify liaison-related information can only emerge when children have memorized a sufficient number of determiner–noun sequences containing an LC. This concept presupposes that there is a high level of interaction between lexical development and morphosyntactic development and that the phonological level is also involved. In effect, the stage 2 schemas (*un*+/nX/, *deux*+/zX/) may regularize the initial segments of the word2 exemplars (*deux nombrils* 'two navels' pronounced *deux zombrils* [døzɔ̃bril]). This aspect of the developmental scenario presupposes that the mobilization of these schemas simultaneously activates units which are represented at the phonological level. As Stemberger (2004) has noted with regard to the possibility of priming regularized forms of the English past tense by means of a vowel situated in the subject noun, this type of interaction requires the simultaneous processing of morphosyntactic and phonological information. If we remember that schemas are connectivity patterns between chunks with phonological content then it becomes clear that their activation simultaneously mobilizes phonological material.

More generally, our developmental conception is strongly compatible with Bybee's ideas (2001), which were developed with adult subjects in mind and which allow us to consider the subsequent development of this system of alternation. In Bybee's model, words, inflected forms and frequent word sequences are memorized in the lexicon and linked together by schemas which encode phonological and semantic generalizations. Certain liaisons are therefore memorized as phonological elements present in these chunks. The establishment of relations between these chunks results in production schemas which may generate liaisons which the subject has never actually heard, and in particular incorrect liaisons which persist, albeit very infrequently, in adults (Desrochers, 1994). Within this perspective, the production of the incorrect liaison *huit-z-arbres* ([ɥizarbr] 'eight trees') consists of activating the specific schema which includes the liaison ([DETERMINER-z-[vowel]-NOUN]$_{\text{plural}}$) rather than the general schema in which it is absent (DETERMINER-NOUN]$_{\text{plural}}$). Moreover, the competition between the two schemas accounts for the variation in some liaisons whose gradual disappearance from the French language seems to take the form of regularization due to the weakening of the least generalized schema.

https://doi.org/10.1017/S0305000908009124 Published online by Cambridge University Press

We would like to emphasize the continuity between the model of acquisition which we have proposed here and the ideas concerning adult performance set out by Bybee. The schemas of the type *un*+/nX/ or *deux*+/zX/, which are characteristic of stage 2, may potentially evolve in various ways during subsequent development. If we accept that the constructions themselves are organized gradually as a network (Kemmer & Barlow, 2000; Tomasello, 2003), then the establishment of relations between different schemas involving a plural determiner and the /z/ liaison (*deux*+/zX/, *trois*+/zX/, *plusieurs*+/zX/, 'two', 'three', 'several') may result in a more abstract construction which associates the /z/ liaison with plural forms (in Bybee's view: [DETERMINER-z-[vowel]-NOUN]$_{\text{plural}}$). To summarize, the constructions in stage 2 of the developmental scenario may evolve towards lexical reorganizations and the encoding of more abstract information, in particular at the morphological level.

In a chapter summarizing studies into the acquisition of alternations in a variety of languages, Bernhardt & Stemberger (1998) commented on our lack of knowledge of this field. They noted a number of errors in the selection of allomorphs which apparently functioned in a similar way to liaison errors (for example, errors in the selection of the *a* and *an* variants of the English indefinite article). The results of the five experiments presented in this article show that liaison is a particularly effective way of exploring this field. In effect, it consists of a phonological alternation situated between two words, the first of which frequently belongs to a closed grammatical class (clitic, determiner). Since liaison errors are indicators of segmentation activities, the acquisition of this alternation makes it possible to observe the way in which children integrate phonological, lexical and morphosyntactic information. The constructionist approach offers a representational format favourable to this type of integration.

REFERENCES

Bates, E. & Goodman, J. C. (1997). On the inseparability of grammar and the lexicon: Evidence from acquisition. *Language and Cognitive Processes* **12**, 507–84.

Bernhardt, B. H. & Stemberger, J. P. (1998). *Handbook of phonological development. From the perspective of constraint-based non linear phonology*. San Diego, London: Academic Press.

Boë, L.-J. & Tubach, J.-P. (1992). *De A à Zut: dictionnaire phonétique du français parlé*. Grenoble: Ellug.

Brooks, P. J. & MacWhinney, B. (2000). Phonological priming in children's picture naming. *Journal of Child Language* **27**, 335–66.

Bybee, J. (2001). Frequency effects on French liaison. In J. Bybee & P. Hopper (eds), *Frequency and the emergence of linguistic structure*, 337–59. Amsterdam, Philadelphia: John Benjamins Publishing Company.

Chevrot, J.-P., Chabanal, D. & Dugua, C. (2007). Pour un modèle de l'acquisition des liaisons basé sur l'usage: trois études de cas. *Journal of French Language Studies* **17**, 103–128.

Chevrot, J.-P. & Fayol, M. (2001). Acquisition of French liaison and related child errors. In M. Almgren, A. Barrena, M. J. Ezeizabarrena, I. Idiazabal & B. MacWhinney (eds),

https://doi.org/10.1017/S0305000908009124 Published online by Cambridge University Press

Research on Child Language Acquisition: Proceedings of the 8th Conference of the International Association for the Study of Child Language, 760–74.
Chevrot, J.-P., Fayol, M. & Laks, B. (2005). La liaison: de la phonologie à la cognition. *Langages* **158**, 3–7.
Chevrot, J.-P., Nardy, A., Barbu, S. & Fayol, M. (2007). Production et jugement des liaisons obligatoires chez des enfants tout-venant et des enfants atteints de troubles du langage: décalages développementaux et différences interindividuelles. *Rééducation orthophonique – Parole(s): aspects perceptifs et moteurs* **229**, 199–220.
Côté, M.-H. (2005). Le statut lexical des consonnes de liaison. *Langages* **158**, 66–78.
Croft, W. & Cruse, D. A. (2004). *Cognitive linguistics*. Cambridge: Cambridge University Press.
De Jong, D. (1994). La sociophonologie de la liaison orléanaise. In C. Lyche (ed.), *French Generative Phonology: Retrospective and Perspectives*, 95–130. Salford: ESRI.
Desrochers, R. (1994). Les liaisons dangereuses: le statut équivoque des erreurs de liaison. *Linguisticae Investicationes* **XVIII:2**, 243–84.
Dugua, C. (2002). *Liaison et segmentation du lexique en français: vers un scénario développemental*. Mémoire de DEA. Grenoble: Université Stendhal.
Dugua, C. (2006). Liaison, segmentation lexicale et schémas syntaxiques entre 2 et 6 ans. Un modèle développemental basé sur l'usage. Thèse de doctorat. Grenoble: Université Stendhal Grenoble3.
Dugua, C., Chevrot, J.-P. & Côte, M.-H. (2003). *Liaison et formation des mots: scénario développemental et conséquences pour le traitement phonologique*. Montpellier: Université Paul-Valéry.
Encrevé, P. (1988). *La liaison avec et sans enchaînement, phonologie tridimensionnelle et usage du français*. Paris: Edition du Seuil.
Fougeron, C., Goldman, J.-P. & Frauenfelder, U. (2001). Liaison and schwa deletion in French: an effect of lexical frequency and competition? *Proceedings of the 7th European Conference on Speech Communication and Technology, Eurospeech 2001*, 639–42.
Goldberg, A. E. (2003). Constructions: a new theoretical approach to language. *Trends in Cognitive Sciences* **7**, 219–24.
Grégoire, A. (1947). *L'apprentissage du langage II – La troisième année et les années suivantes*. Genève: E. Droz.
Judd, C. M., McLelland, G. & Culhane, S. E. (1995). Data analysis: continuing issues. The everyday analysis of psychological data. *Annual review of psychology* **46**, 433–65.
Kemmer, S. & Barlow, M. (2000). Introduction: A usage-based conception of language. In M. Barlow & S. Kemmer (eds), *Usage-based models of language use*, VII–XXVIII. Stanford California: CSLI Publications.
Malécot, A. (1975). French liaison as a function of grammatical, phonetic and paralinguistic variables. *Phonetica* **32**, 161–79.
Mattys, S. L. & Jusczyk, P. W. (2001). Do infants segment words or recurring continuous patterns? *Journal of Experimental Psychology: Human Perception and Performance* **27**, 644–55.
Moisset, C. (2000). Variable liaison in parisian French. Unpublished doctoral dissertation. University of Pennsylvania.
Morel, E. (1994). Le traitement de la liaison chez l'enfant: études expérimentales. *TRANEL* **21**, 85–95.
Morin, Y.-C. (2003 [1998]). Remarks on prenominal liaison consonant in French. In S. Ploch (ed.), *Living on the Edge – 28 Papers in Honour of Jonathan Kaye*, 385–400. Berlin: Mouton de Gruyter.
Nardy, A. (2003). *Production et jugement d'acceptabilité entre 2 et 6 ans: Aspects psycholinguistiques et sociolinguistiques de l'acquisition des liaisons*. Mémoire de DEA. Grenoble: Université Stendhal.
New, B., Pallier, C., Ferrand, L. & Matos, R. (2001) Une base de données lexicales du français contemporain sur internet: LEXIQUE, *L'Année Psychologique* **101**, 447–62 (www.lexique.org).

https://doi.org/10.1017/S0305000908009124 Published online by Cambridge University Press

Peters, A. M. (1985). Language segmentation: operating principles for the perception and analysis of language. In D. Slobin (ed.), *The crosslinguistic study of language acquisition*, 1029–67. Hillsdale NJ: Erlbaum.
Pine, J. & Lieven, E. (1993). Reanalysing rote-learned phrases: Individual differences in the transition to multi word speech. *Journal of Child Language* **20**, 551–71.
Pine, J. & Lieven, E. (1997). Slot and frame patterns in the development of the determiner category. *Applied psycholinguistics* **18**, 123–38.
Schane, S. A. (1968). *French phonology and morphology*. Cambridge, MA: MIT Press.
Spinelli, E., McQueen, J. M. & Cutler, A. (2003). Processing resyllabified words in French. *Journal of Memory and Language* **48**, 233–54.
Spinelli, E. & Meunier, F. (2005). Le traitement cognitif de la liaison dans la reconnaissance de la parole enchaînée. *Langages* **158**, 79–88.
Stemberger, J. P. (1990). Wordshape errors in language production. *Cognition* **35**, 123–57.
Stemberger, J. P. (2004). Phonological priming and irregular past. *Journal of Memory and Language* **50**, 82–95.
Ternes, E. (1977). Konsonantische Anlautveränderungen in den keltischen und romanischen Sprachen. *Romanistisches Jahrbuch* **28**, 19–53.
Tomasello, M. (2003). *Constructing a language: a usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.
Tranel, B. (2000). Aspects de la phonologie du français et la théorie de l'optimalité. *Langue française* **126**, 39–72.
Vinter, S. (2001). Les habiletés phonologiques chez l'enfant de deux ans. *GLOSSA* **77**, 4–19.
Wauquier-Gravelines, S. (2003). Du réalisme des formalisations phonologiques contemporaines: que nous apprennent les données d'acquisition? In J. P. Angoujard & S. Wauquier-Gravelines (eds), *Phonologie: champs et perspectives*, 9–34. Paris: Presses de l'Ecole Normale Supérieure de Fontenay-St Cloud.
Wauquier-Gravelines, S. & Braud, V. (2005). Proto-déterminant et acquisition de la liaison obligatoire en français. *Langages* **158**, 53–65.
Yersin-Besson, C. & Grosjean, F. (1996). L'effet de l'enchaînement sur la reconnaissance des mots dans la parole continue. *L'Année psychologique* **96**, 9–30.

APPENDIX

Material used in experiments 1, 2, 4 and 5.

Frequencies (occurrences per million) are given by the French database Lexique (New, Pallier, Ferrand & Matos, 2001). The plural orientation of the vowel initial word2s was obtained by the judgments of seventy native French-speaking adults who chose the more familiar between two sequences (e.g. *un ours* 'one/a bear' – *des ours* 'some bears'). For each noun, we established a ratio that gave the tendency for the noun to be plural: number of plural choices/(number of plural + number of singular choices). A ratio close to 1 suggests that a noun is more often employed in the plural form and is more likely to be preceded by the /z/ liaison. In contrast, a ratio close to 0 suggests a frequent singular use of the noun. Chevrot, Chabanal & Dugua (2007) have shown that this judgment-based ratio correlates with the assessment of the plural orientation using a French database.

https://doi.org/10.1017/S0305000908009124 Published online by Cambridge University Press

Determiners: experiments 1, 4 and 5	Frequency
un 'a/one'	13 617
deux 'two'	1558
Consonant-initial nouns: experiments 4 and 5	
lavabo 'washbasin'	17
nombril 'navel'	5
nuage 'cloud'	77
zèbre 'zebra'	5

Vowel-initial nouns: experiments 1 and 2	Frequency	Ratio towards the plural
ours 'bear'	17	0·307
arbre 'tree'	208	0·604
élephant 'elephant'	15	0·293
écureuil 'squirrel'	10	0·254
âne 'donkey'	19	0·137

https://doi.org/10.1017/S0305000908009124 Published online by Cambridge University Press