

BUCLD 35 Proceedings
To be published in 2011 by Cascadilla Press
Rights forms signed by all authors

Processing of French Liaisons in Toddlers

Mireille Babineau and Rushen Shi

1. Introduction

Liaison is a common phenomenon in French. It is manifested by the surfacing of an underlying consonant between two words. The consonant is the underlying coda of the first word which only surfaces when the second word is vowel-initial, and this consonant is phonetically the onset consonant of the second word (e.g., *les amis* is phonetically [le zami], “the friends”). The first word is usually one of a set of frequently occurring words, mainly function words. The underlying coda does not surface in other contexts, neither when preceding a consonant-initial second word, nor when occurring with no following word. For example, the underlying coda /z/ of *les* (“the”) is not phonetically realized when it co-occurs with the word *taux* (“rates”, [le to]). The surfacing of the underlying coda as the onset consonant of the subsequent vowel-initial word (e.g., *les amis*, [le zami]) satisfies the phonological constraint that favors consonant-initial onset for syllable (Tranel, 1996). However, the resyllabification results in a mis-alignment between the word onset and syllabic onset.

In the acquisition literature, vowel-initial word segmentation has been shown to be harder than consonant-initial word segmentation, and this difficulty is related to resyllabification. For example, the co-occurrence of *cold* and *ice* generally causes the coda /d/ of *cold* to be resyllabified as the onset consonant of *ice*, making *ice* difficult to parse. Whereas the segmentation of consonant-initial words starts from 6 months of age (e.g., Bortfeld, Morgan, Golinkoff, & Rathbun, 2005; Jusczyk & Aslin, 1995), vowel-initial word segmentation starts much later. In Seidl and Johnson (2008) infants aged 11 months segmented vowel-initial words only when they occurred in a salient position, such as the utterance-initial position (e.g. *Eff runs a circus in Toronto*). Mattys and Jusczyk (2001) showed that when vowel-initial words (e.g., *ice*, *ash*) were resyllabified with the coda consonant of a preceding word (e.g., *Weird ice no longer surprises anyone*), only 16-month-olds, but not 8.5-, 10.5- or 13-month-olds, succeeded in segmenting them. Interestingly, in another experiment they found that even 8.5-

* Mireille Babineau and Rushen Shi, Université du Québec à Montréal, babineau.mireille@courrier.uqam.ca. This research is supported by SSHRC, NSERC and CFI grants to the second author. We are grateful to Andréane Melançon, Marilyn Cyr, and Amy Arnott for research assistance. We also thank all the parents and infants for their participation.

month-olds did not mis-segment such resyllabified cases (e.g. *Weird ice no longer surprises anyone*) as containing a consonant-initial word (e.g., *dice*), whereas they did segment *dice* from sentences that truly contained this consonant-initial word (e.g., *Two dice can be rolled without difficulty*). This suggests that inherently consonant-initial words and vowel-initial words that are resyllabified (e.g., *two dice* versus *weird ice*) are not entirely ambiguous phonetically. Nevertheless, the phonetic differences were not strong enough to enable the younger infants to segment the vowel-initial words from the resyllabified environment. Liaison is unlike such general resyllabification cases. It is a phonological process that involves complete resyllabification. Thus, the segmentation of vowel-initial words in liaison cases is likely harder. This issue has not been examined in any perceptual experiment with infants.

Another likely source of difficulty for word segmentation in liaison concerns the encoding status of the first word. For non-liaison-related words (e.g., *cold*, *that*, *chaque* “each”), their final consonant is present in all contexts. For instance, the coda /t/ of *that* is surfaced before a vowel-initial word *ice* (resyllabified with *ice*), in isolation or utterance positions (e.g., *see that*), and before a consonant-initial word (e.g., *that chair*). Crucially, in the two latter types of environment (i.e., not preceding a vowel-initial word), /t/ is phonetically realized as the coda of *that*, allowing the learner to encode it as a part of the word *that*. The encoded form *that* may in turn assist the parsing of the vowel-initial word that follows it, e.g., *that ice*. There is evidence that infants indeed encode the forms of frequent words including function words starting from 6 months of age and use them to segment the adjacent consonant-initial word (Bortfeld et al., 2005; Hallé, Durand, & de Boysson-Bardies, 2008; Shi, Cutler, Werker, & Cruickshank, 2006; Shi & Lepage, 2008; Shi, Werker, & Cutler, 2006). The coda status in liaison-causing words (i.e., the first word in a liaison case) is more complex than that of non-liaison-causing words. The underlying coda of liaison-causing words (the first word) such as *les* and *un* only surfaces when a vowel-initial word follows it (e.g., /z/ in *les amis* [le zami]), but is mute in all other contexts (e.g., *les taux* [le to], *les* [le] in citation form or in the utterance final position). The learner receives no direct evidence that such words (e.g., *les*, *un*) contain a coda consonant. Therefore, finding the onset of vowel-initial words in liaison (i.e., the second word) is likely more difficult than in general resyllabification cases.

A possible solution for segmenting vowel-initial words in liaison may be provided by statistical cues. If a vowel-initial word occurs in various kinds of liaisons (e.g., *les* [z]ami, *un* [n]ami, *petit* [t]ami), the transitional probability between the vowel-initial word and the preceding consonant is lowered, providing a plausible basis for segmenting the vowel-initial word. Previous studies showed that infants from 7 months of age begin to use statistical cues to find word boundaries (e.g., Pelucchi, Hay & Saffran, 2009; Saffran, Aslin, & Newport, 1996; Thiessen & Saffran, 2003). However, in those segmentation studies, statistical computations were always aligned with syllabic boundaries. That is, the unit of statistical computations in those segmentation studies was the

syllable. It is unknown if infants are able to segment subsyllabic units based on their statistical distribution with adjacent phonetic elements. Liaison involves subsyllabic word units, thus offering a unique chance for studying this question.

The statistical cues in variable liaison cases indicate the beginning of vowel-initial words. But this parsing is misaligned with the syllabic boundary (e.g., *amis* in *les amis* [le zami]), an outcome that is disfavored by the syllabicity constraint proposed by several researchers. For example, in linguistics, phonologists consider the existence of a syllabic constraint as part of the grammar, which favors the alignment of the syllabic edges with the word edges (e.g., McCarthy & Prince, 1994). Norris and colleagues (Norris, McQueen, Cutler, & Butterfield, 1997) proposed that word segmentation is guided by the universal constraint of syllabicity, which does not leave a consonant behind. English-speaking adults in their experiments resisted segmenting subsyllabic vowel-initial words that left behind its onset consonant (e.g. *apple* from *fapple*). In subsequent infant perceptual experiments, they found the same pattern of results with English-hearing 12-month-olds (Johnson, Jusczyk, Cutler, & Norris, 2003). Experiments in French (Content, Dumay, & Freuenfelder, 2000) showed that adult listeners' word segmentation respected the syllable boundaries intended by the speaker; that is, syllabic boundaries were important for perceiving word onset. The syllabicity constraint would hence predict that segmentation of vowel-initial words in liaison is hard because it violates the integrity of the syllable. Children's production errors of vowel-initial words in liaison cases seem to suggest that the consonant-initial surface forms are stored in memory (e.g. *des namis* instead of *des [z]amis*; Dugua, Chevrot & Fayol, 2006). These errors can be linked to the syllabicity perceptual bias, despite statistical support for vowel-initial word segmentation in variable liaison cases that infants may have heard in the input. Infants may have just retained and stored one of the surface forms, such as *nami*, as the preferred form. However, the errors could also be because the children in Dugua, et al. (2006) had never previously heard variable liaison cases for particular words (e.g., always hearing *un [n]ami*), thus lacking statistical support for vowel-initial parsing.

Therefore, it is not known whether infants can use statistical cues to segment vowel-initial words in liaison, or whether their segmentation is strictly driven by the syllabicity constraint. To study this issue, it is necessary to conduct experiments that put subsyllabic statistical cues and syllabicity in conflict. A few infant studies have previously examined conflicting cues to word segmentation. Thiessen and Saffran (2003) showed that English-learning 7-month-olds weighed statistical cues more than stress cues for word segmentation when the two types of cues were in conflict (i.e., indicating different word boundaries). By 9 months of age, infants showed more reliance on stress for locating word boundaries. Johnson and Jusczyk (2001) also showed that English-learning 8-month-olds relied on stress more than statistics for word segmentation. Moreover, they found that when statistical cues were in conflict with coarticulation cues, 8-month-olds relied on coarticulation cues to find word boundaries. These studies seem to suggest that younger infants use statistical

cues, but when they get older, they discover and use other types of cues based on their increased experience with the speech input. It appears that statistical computation is a primary mechanism, since it is available at such an early age. However, these studies involved the perfect alignment of syllable edges with statistical units. No study to date has tested if infants can do statistical word segmentation for subsyllabic units, or if there is a language-general syllable-alignment bias present at an early age to constrain subsyllabic statistical parsing. Our study aimed at addressing these questions, using French liaison as the testing case.

In previous preferential looking experiments Quebec-French-learning infants segmented monosyllabic words (Marquis & Shi, 2008; Shi & Lepage, 2008) and bisyllabic words (Polka & Sundara, 2003) before one year of age. In experiments with European French, infants showed the ability to segment bisyllabic words using between-syllable distributional patterns at 16 months of age, whereas 12-month-olds used the strategy of segmenting each syllable as an individual word (Nazzi, Iakimova, Bertoncini, Frédonie, & Alcantara, 2006). Bisyllabic word segmentation was later observed in 12-month-old French infants in a more sensitive ERP experiment (Goyet, de Schonen & Nazzi, 2010). None of these French studies was designed to test syllabic versus subsyllabic word segmentation. In the present study we pitted subsyllabic statistical cues against syllabic integrity to determine the early segmentation biases and mechanisms in infants.

2. Experiment 1

2.1.1. Participants and Stimuli

Participants were monolingual Quebec-French-learning infants aged 20 months. Stimuli were 8 sentences containing the pseudo-noun *onche* and 8 containing the pseudo-noun *èque* in four liaison contexts (/z/, /n/, /t/, /r/). The sentences contained infrequent content words. One common adjective or function word preceded the vowel-initial pseudo-noun in each sentence. Two sentences for each of the four liaison types were constructed and the vowel-initial pseudo-noun was preceded in each of these sentences by a different word (e.g., for the /z/ liaison type, the pseudo-noun *onches* was preceded by *gros* and *ces*, i.e., *gros [z]onches* and *ces [z]onches*). In addition, stimuli included isolated vowel-initial pseudo-nouns (*onche* and *èque*).

A female native speaker of Quebec French recorded multiple tokens of the sentences and the pseudo-nouns in isolation in infant-directed speech style. The recording was done in an IAC acoustic chamber, at 44.1 Khz sampling frequency with a bit rate of 16 bits. The final familiarization stimuli consisted of 2 tokens of each of the 16 sentences. The final test stimuli consisted of 24 isolated productions of each the pseudo-nouns (12 *onche* and 12 *èque*).

The visual image for all trials was an abstract picture, presented synchronously with the auditory stimuli during each trial of the experiment. The experiment started with a pre-trial, presenting sounds of water bubbles with the

same abstract picture, which served to acquaint the infant to the procedure. The attention-getter between trials was a video of a moving star with sounds of a cricket and was designed to attract infants' attention between trials.

2.1.2. Procedure

Infants were individually tested in a visual preferential procedure in a sound-attenuated room. The auditory stimuli were presented from loudspeakers that were placed together with a central display monitor. The infant sat on the parent's lap in front of the monitor, which presented the visual stimuli. The parent heard masking music through headphones. The task was run with the Habit program (Cohen, Atkinson, & Chaput, 2000). The experiment advanced automatically to the test phase after the familiarization phase was completed.

A researcher blind to the audio-visual materials observed the infant's eye movement from an adjacent room through a closed circuit TV. Each trial was initiated when the infant looked toward the monitor. The researcher kept a computer key pressed down whenever the infant looked at the monitor. Each trial terminated automatically once the maximum trial length was reached. All experimental sessions were videotaped and then coded offline frame by frame (30 frames per second) by another blind researcher.

2.1.3. Design and predictions

The experiment consisted of a familiarization phase and a test phase. One group of infants was familiarized with sentences containing the pseudo-noun *onche*, and another group with sentences containing the pseudo-noun *èque* (Table 1). During the test phase all infants heard isolated tokens of *onche* and *èque*, in alternating trials. One test trial type was "familiarized", and the other was "non-familiarized". Specifically, for infants who were familiarized with sentences containing *onche*, the test trials presenting *onche* in isolation were "familiarized", and those presenting isolated *èque* tokens were "non-familiarized". For the other group, who were familiarized with *èque* sentences, the test trials presenting isolated *èque* tokens were "familiarized", and those presenting isolated *onche* tokens were "non-familiarized". The test phase began with either the *onche* or *èque* trial, counter-balanced across infants.

The length of each familiarization trial was 6.4 sec, and the length of each test trial was 18.5 sec. Infants in each familiarization group heard 8 familiarization trials presenting sentences containing one of the two target pseudo-nouns (*onche* or *èque*). Each familiarization trial contained 2 sentences, and the average sentence duration was 2.39 sec. The average inter-sentence interval in each trial was 998 msec. In the test phase, 5 trials were presented for each type ("familiarized" versus "non-familiarized") for a total of 10 test trials. Each test trial contained 12 isolated tokens of either the familiarized word or the non-familiarized word. The average duration for isolated word forms was 573 msec. The average inter-stimulus interval in each test trial was 1001 msec.

In the familiarization phase, statistical cues supported the correct word boundary because various underlying liaison codas surfaced with the same vowel-initial word (*onche* or *èque*). The transitional probability between the vowel-initial target (*onche* or *èque*) and the preceding elements is low. However, syllabicity supports a consonant-initial interpretation, which is in conflict with the statistical cues that support a vowel-initial interpretation. We predicted that if 20-month-olds can use statistical cues to segment vowel-initial words in liaison, they should then be able to discriminate between the familiarized and non-familiarized test trials.

Table 1. Stimuli and design for the Experiment 1.

Group 1 Familiarization	Group 2 Familiarization
<i>J'ai trouvé mon onche sur le pavillon.</i>	<i>Il y avait mon èque sur l'étagère.</i>
<i>Il y avait couramment un onche.</i>	<i>On ne réussit guère un èque.</i>
<i>C'était un petit onche qui était ravi.</i>	<i>Voici un petit èque que je trouve fou.</i>
<i>Je suis camouflé au grand onche.</i>	<i>Je t'ai fréquenté au grand èque.</i>
<i>Ça n'englobe pas ces onches.</i>	<i>J'avais déniché ces èques.</i>
<i>Ces gros onches sont de sales voyous.</i>	<i>Ces gros èques le chagrinent beaucoup.</i>
<i>Ce premier onche veut s'enfuir.</i>	<i>Le premier èque va rugir.</i>
<i>Il a discuté du dernier onche.</i>	<i>Voilà la plainte du dernier èque.</i>
Test trials: <i>onche (familiarized) versus èque (non-familiarized)</i>	Test trials: <i>onche (non-familiarized) versus èque (familiarized)</i>

2.2. Results

Looking times for the two test trial types (“familiarized” and “non-familiarized”) were analyzed in a paired t-test. As shown by the left two columns of Figure 1, infants showed no looking time difference while listening to the familiarized vowel-initial pseudo-noun versus to the non-familiarized vowel-initial pseudo-noun, $t(15) = 0.29$, $p = .776$. This suggests that infants did not have a vowel-initial interpretation, thus they did not rely on the statistical cues in the familiarization stimuli to segment the vowel-initial word in liaison cases. That is, when statistical cues and syllabicity were in conflict, 20-month-olds show no evidence of being able to use the statistical cues to parse subsyllabic units. We can hypothesize that this failure in vowel-initial word segmentation may be linked to a strong bias for syllabic integrity. Experiment 2 tested this bias by familiarizing 20-month-olds with the same vowel-initial pseudo-nouns in variable liaison cases but testing them with isolated consonant-initial words (i.e., syllable-alignment interpretation).

3. Experiment 2

3.1. Participants, Stimuli, Design and Predictions

Participants were another group of 20-month-old monolingual Quebec-French-learning infants. Familiarization stimuli were the same as in Experiment 1. The test tokens were isolated pseudo-words *zonches* and *zèques*. We decided to use /z/ as the onset consonant of the test words for several reasons. It is one of the most common liaison consonants in French, because it is the underlying coda of highly frequent function words such as *les* (“the”, a definite article), *des* (“some”, an indefinite article), etc. The coda *z* is also more grammatically involved as a plural marker, even with vowel-initial verbs (e.g., *ils [z]ont*, “they have”). In the natural speech input of 20-month-olds, the liaison consonant /z/ has likely occurred with a great number of vowel-initial content words following it. That is, the statistical cues for /z/ to belong to the first word (e.g., *les*, *des*) should be strong. Furthermore, /z/ is less related lexically to consonant-initial words in general, since z-initial words in French like *zéro* (zero) or *zèbre* (zebra) are a very small set. Therefore, if a segmentation effect is found with /z/, it would be a strong support for the syllable-alignment constraint.

The same female native speaker of Quebec French as Experiment 1 recorded multiple productions of the consonant-initial pseudo-nouns *zonches* and *zèques*. All recording aspects were identical to those of Experiment 1. The final test stimuli consisted of 12 isolated productions of each pseudo-noun for a total of 24 tokens (2 types x 12 tokens).

Table 2. Stimuli and design for Experiment 2.

Group 1 Familiarization (identical to Exp 1)	Group 2 Familiarization (identical to Exp 1)
<i>J'ai trouvé mon onche sur le pavillon.</i>	<i>Il y avait mon èque sur l'étagère.</i>
<i>Il y avait couramment un onche.</i>	<i>On ne réussit guère un èque.</i>
<i>C'était un petit onche qui était ravi.</i>	<i>Voici un petit èque que je trouve fou.</i>
<i>Je suis camouflé au grand onche.</i>	<i>Je t'ai fréquenté au grand èque.</i>
<i>Ça n'englobe pas ces onches.</i>	<i>J'avais déniché ces èques.</i>
<i>Ces gros onches sont de sales voyous.</i>	<i>Ces gros èques le chagrinent beaucoup.</i>
<i>Ce premier onche veut s'enfuir.</i>	<i>Le premier èque va rugir.</i>
<i>Il a discuté du dernier onche.</i>	<i>Voilà la plainte du dernier èque.</i>
Test trials: <i>zonches (familiarized) versus zèques (non-familiarized)</i>	Test trials: <i>zonches (non-familiarized) versus zèques (familiarized)</i>

The procedure and design were identical to those of Experiment 1, except that the test words were *zonches* and *zèques* (see Table 2). In fact, the test words *zonches* and *zèques* matched the /z/ liaison surfaced forms in the familiarization sentences (*ces [z]onches*, *gros [z]onches*, *ces [z]èques*, *gros [z]èques*), although these test words were not the correct segmentation according to the adult grammar. The stimuli in the test trials were consistent with a consonant-initial interpretation based on the syllable-alignment bias. If infants were driven by this bias, they should distinguish the two test trial types (“familiarized” versus “non-familiarized”) and show a systematic preference.

3.2. Results

Infants looked longer during the test trials containing the non-familiarized pseudo-noun (see the middle of Figure 1). Looking times to the familiarized consonant-initial form versus the non-familiarized consonant-initial form were significantly different ($t(12) = -2.43, p = .032$), suggesting that infants interpreted the second word in the liaison cases in the familiarization stimuli as consonant-initial, a syllable-alignment bias. These results contrast with those of Experiment 1, which showed no vowel-initial segmentation, i.e., no evidence of subsyllabic statistical segmentation. This was confirmed by analyzing the data of both experiments in a 2x2 ANOVA, with Test Trial Type (“familiarized” versus “non-familiarized”) as the within-subject factor and Experiment (Experiment 1 subsyllabic versus Experiment 2 syllabic) as the between-subject factor. No main effect of Test Trial Type ($F(1, 27) = 2.627, p = .117$) nor Experiment ($F(1, 27) = 0.223, p = .641$) was found, but the Test Trial Type x Experiment interaction was nearly significant ($F(1, 27) = 4.033, p = .055$).

Taking both experiments together, we may interpret that infants at 20 months failed to show evidence of statistical based segmentation of subsyllabic units in liaison cases. Instead, they relied on the syllable edges to find word boundaries, i.e., a strong bias for syllabic integrity at this age. The syllabicity bias led to the misperception shown in Experiment 2 (e.g., perceiving *onche* as *zonches*) despite the variable liaison cues for the vowel-initial interpretation.

Although the segmentation bias for syllabic integrity appears strong, infants acquiring French must ultimately learn to segment subsyllabic units such as vowel-initial words in liaison cases. The goal of Experiment 3 was to test if older infants can successfully segment vowel-initial words based on statistical cues available in variable liaison contexts. That is, we examined their ability to perform subsyllabic statistical parsing against conflicting syllable alignment.

4. Experiment 3

4.1. Participants, Stimuli, Design and Predictions

Participants were 24-month-old monolingual Quebec-French-learning infants. The stimuli, procedure and design were identical to those of Experiment 1 (see Table 1). As in Experiments 1 and 2, the statistical cues in the variable liaison cases presented in familiarization sentences supported vowel-initial interpretation, but the consonant-initial realizations in liaison supported the syllable-alignment constraint. The two cues are thus conflicting. The test phase which presented vowel-initial forms (*onche* vs *èque*) aimed at testing 24-month-olds’ subsyllabic statistical parsing ability. Our predictions were that if by 24 months of age infants still have a syllable-alignment bias, like the 20-month-olds, then no discrimination of the two test trial types (“familiarized” versus “non-familiarized”) should be observed. On the other hand, if they have learned to use the subsyllabic statistical cues to parse the vowel-initial word in liaison cases, a systematic discrimination response should be found.

4.2. Results

As shown by the rightmost two columns of Figure 1, infants significantly discriminated the familiarized vowel-initial pseudo-noun from the non-familiarized vowel-initial pseudo-noun, $t(14) = 3.798$, $p = .002$. This result suggests that by 24 months of age infants have learned to counter the syllable-alignment bias with subsyllabic statistical segmentation, and this ability led to a vowel-initial interpretation corresponding to the French liaison rule. These results contrast with those of Experiment 1. In that experiment 20-month-olds did not show a discrimination of the same test stimuli. That is, there was no evidence of the ability to perform subsyllabic statistical parsing against the syllable-alignment bias. By 24 months of age, infants have acquired this ability.

This difference in the results between Experiment 1 and 3 was confirmed by a 2x2 mixed ANOVA with Test Trial Type (“familiarized” versus “non-familiarized”) as the within-subject factor and Age (20 months versus 24 months) as the between-subject factor. No main effect of Age was obtained ($F(1, 29) = .561$, $p = .46$). A significant Test Trial Type x Age interaction ($F(1, 29) = 6.051$, $p = .02$) was observed. There was also a significant main effect of Test Trial Type ($F(1, 29) = 8.249$, $p = .008$), exhibiting an overall longer looking time for the familiarized trials than non-familiarized trials. This main effect of Test Trial Type was due to the highly significant looking difference shown by the 24-month-olds. Indeed, the 20-month-olds in Experiment 1 showed no looking time difference for familiarized vs. non-familiarized words.

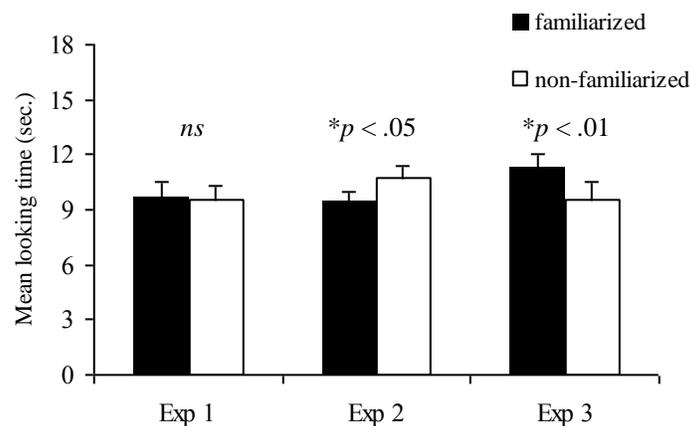


Figure 1. Infants’ looking (listening) times to the two types of test trials, familiarized versus non-familiarized. The test words in Experiment 1 and 3 (*onche*, *èque*) represented the statistical parse, and those in Experiment 2 (*zonche*, *zèque*) represented the syllable-alignment parse.

5. Discussion

In this study we hypothesized that statistically-based word segmentation may be subject to a certain limit. In particular, we asked whether a general bias for preserving syllabic integrity exists in early infancy, constraining the possible types of statistical segmentation. Our results showed that infants failed to use statistical cues to segment vowel-initial words in liaison even at 20 months of age. This contrasts with the remarkable ability of infants around 7 and 8 months of age to segment consonant-initial words using statistics, as shown by previous studies (e.g., Saffran, Aslin, & Newport, 1996; Thiessen & Saffran, 2003), in which syllable integrity was preserved. Only at 24 months of age did our infants succeed in using subsyllabic statistical cues to segment vowel-initial words in liaison. Importantly, our 20-month-olds misperceived the vowel-initial words as being consonant-initial due to the syllable integrity bias, despite the statistical cues supporting vowel-initial parsing. Therefore, French-hearing infants require a significant amount of learning in order to counter the general syllabicity bias for successful segmentation of subsyllabic word units in liaison.

The consonant-initial parsing in our younger infants is consistent with the syllable-alignment constraint proposed in the frameworks of generative linguistics (e.g., McCarthy & Prince, 1994) and psycholinguistics (Norris, McQueen, Cutler, & Butterfield, 1997; Content, Dumay, & Freuenfelder, 2000). Our findings suggest that early in acquisition the syllable-alignment bias guides word parsing and strongly constrains certain kinds of statistical segmentation. As a consequence, subsyllabic word segmentation is disfavored despite statistical support, whereas syllable-aligned units would enhance statistical segmentation. Consistent with the latter case, previous studies that demonstrated statistical word segmentation in infants as young as 7 to 8 months of age indeed involved statistical units that were aligned with the syllable (e.g. Saffran, Aslin & Newport, 1996; Thiessen & Saffran, 2003). Therefore, complying with the syllable-alignment bias yields optimal word segmentation. However, violations of the bias lead to inhibition of statistics-based word segmentation, as shown in the present study. These outcomes are reasonable since a plausible and efficient learning system should favor the more common processing units for natural languages, and the system should be constrained against too many kinds of computational units including the uncommon or irrelevant ones.

Ultimately, subsyllabic processing is needed for vowel-initial parsing in liaison cases. Experiment 3 shows that by 24 months of age infants have learned to do so by using statistical cues, although the boundary supported by the cues is mis-aligned with the syllabic edge. We are currently testing 24-month-olds on the stimuli of Experiment 2 (syllable-alignment parsing), and the preliminary results suggest that infants of this age no longer operate with a syllable-aligned interpretation when faced with conflicting statistical cues, unlike our 20-month-old infants. The older infants appeared to have learned to overcome the syllable-alignment constraint for parsing liaison cases.

It is interesting to note that Experiments 2 and 3 show opposite directions of looking preference (novelty preference versus familiarity preference). Different directions of preference have been reported and discussed by a number of researchers (e.g. Hunters & Ames, 1988; Thiessen & Saffran, 2003), who examined several factors underlying the difference. Prolonged familiarization of the same stimuli usually leads to a novelty preference during the test phase. Another factor is the level of difficulty of the task. An easy task is likely to produce a novelty preference, whereas a difficult task is likely to yield a familiarity preference. Our results can thus be interpreted as follows: The syllable-alignment bias which guided 20-month-olds in their consonant-initial interpretation is a strong ability for infants of this age, thus yielding a novelty preference. As mentioned earlier in this article, consonant-initial word segmentation starts at a very early age and is thus an easy process for 20-month-old infants. On the other hand, the vowel-initial segmentation in liaison by 24-month-olds is a newly acquired ability. The computation of subsyllabic statistics seems to be more demanding than the processing of the syllable as the unit, which may explain the familiarity preference in this older age group. The novelty preference in our 20-month-olds is particularly impressive since the familiarization sentences presented to each infant contained only four occurrences of the surface form that matched one of the test words. This suggests that the novelty preference was not resulted from prolonged familiarization, but rather, it reflected a robust syllable-alignment bias.

The segmentation of vowel-initial words in usual resyllabification cases in English emerges at around 16 months of age (Mattys & Jusczyk, 2001), months earlier than the liaison-related vowel-initial word segmentation observed in our study. Even when younger infants failed to segment vowel-initial words, their perception of usual resyllabification and liaisons differ. In particular, English-learning 8.5-month-olds in Mattys and Jusczyk (2001) did not have a consonant-initial interpretation (e.g. *dice*) after hearing a vowel-initial word in a resyllabification context (e.g. *cold ice*). In contrast, our French-learning infants misperceived vowel-initial words in liaison cases as consonant-initial (e.g., *zonches* from *ces onches*) even at 20 months of age. These differences are reasonable given the distinct nature of the type of resyllabification in Mattys and Jusczyk (2001) as opposed to liaison resyllabification. As discussed in the Introduction, usual resyllabifications are often not completely resyllabified (e.g. *cold ice*), leaving acoustical cues to potentially help un-do the resyllabification. Liaison, on the other hand, produces phonetically complete resyllabifications (e.g. *les amis* [le zami]), making the vowel-initial word hard to parse. In addition to the different degrees of resyllabification, the coda consonant of non-liaison-causing words is often present as a coda in non-resyllabified contexts (e.g., *cold room*), allowing the infant to encode its full form and to use it for word segmentation in resyllabified cases. Liaison is disadvantaged in this aspect. The underlying coda of the liaison-causing words is phonetically realized only when such words precede a vowel-initial word, in which case the coda surfaces as the onset consonant of the vowel-initial word. Otherwise, the coda is always mute

(e.g. *les taux* [le to]). Thus, the surfaced consonant in liaison is more likely to be misperceived as the onset consonant of the following vowel-initial word.

In French, there is an equivalent of the English resyllabification, which occurs widely in the language (e.g. *chaque ours*, “each bear”; *chaque* is a regular CVC, non-liaison-causing word). The resyllabification in these cases is arguably not entirely complete. Like the comparable cases in English, the coda of *chaque* is realized in all contexts, and is often not resyllabified with the following word (e.g., *chaque livre*, “each book”). It is thus possible that vowel-initial segmentation in usual resyllabification cases (*chaque ours*) emerges at a younger age than in liaison cases. In future studies we plan to test whether vowel-initial words in usual resyllabifications in French are easier for infants to segment, and if statistical cues are easier to use in such cases.

In sum, in three experiments we showed that when subsyllabic statistical cues and syllable-alignment cues were in conflict (i.e., indicating different word boundaries), French-learning 20-month-olds’ word segmentation was guided by a general syllable-alignment constraint. Close to 24 months of age infants overcame the syllable-alignment constraint, and learned to use subsyllabic statistical cues to successfully segment vowel-initial words in liaison contexts.

References

- Bortfeld, H., Morgan, J. L., Golinkoff, R. M., & Rathbun, K. (2005). Mommy and me: Familiar names help launch babies into speech-stream segmentation. *Psychological Science, 16*, 298-304.
- Cohen, L. B., Atkinson, D. J., & Chaput, H. H. (2000). Habit 2000: A new program for testing infant perception and cognition. (Version 2.2.5c) [Computer software]. Austin: University of Texas.
- Content, A., Dumay, N., & Frauenfelder, U. H. (2000). The role of syllable structure in lexical segmentation in French: Helping listeners avoid mondegreens. In A. Culter, J. M. McQueen & R. Zondervan (Ed.), *Spoken word access processes* (pp.39-42). Nijmegen, The Netherlands: Max-Planck Institute for Psycholinguistics.
- Dugua, C., Chevrot, J.-P. & Fayol, M. (2006). Liaison, segmentation des mots et schémas syntaxiques entre 2 et 6 ans : un scénario développemental. In D.-J. Duché (Ed.), *Entretiens de Bichat : Orthophonie* (pp. 230-244). Paris : Expansion, formation et éditions.
- Goyet, L., de Schonen, S., & Nazzi, T. (2010). Words and syllables in fluent speech segmentation by French-learning infants: an ERP study. *Brain Research, 1332*, 75-89.
- Hallé, P., Durand, C., & de Boysson-Bardies, B. (2008). Do 11-month-old French infants process articles? *Language and Speech, 51*, 23–44.
- Hunter, M. A., Ames, E. W. (1988). A multifactor model of infant preferences for novel and familiar stimuli. In C. Rovee-Collier, & L. P. Lipsitt (Eds.), *Advances in Infancy Research* (pp. 69–95). Norwood, NJ: Ablex.
- Johnson, E. K., & Jusczyk, P.W. (2001). Word segmentation by 8-month-olds: When speech cues count more than statistics. *Journal of Memory and Language, 44*, 548-567.

- Johnson, E. K., Jusczyk, P. W., Cutler, A., & Norris, D. (2003). Lexical viability constraints on speech segmentation by infants without a lexicon. *Cognitive Psychology*, *46*, 65-97.
- Jusczyk, P. W., & Aslin, R. N. (1995). Infants' detection of the sound patterns of words in fluent speech. *Cognitive Psychology*, *29*, 1-23.
- Marquis, A., & Shi, R. (2008). Segmentation of verb forms in preverbal infants. *Journal of the Acoustical Society of America*, *123* (4), EL105-EL110.
- Mattys, S. L., & Jusczyk, P. W. (2001). Do infants segment words or recurring contiguous patterns? *Journal of Experimental Psychology: Human Perception and Performance*, *27* (3), 644-655.
- McCarthy, J. J., & Prince, A. (1994). The emergence of the unmarked: Optimality in prosodic morphology. In M. González (Ed.), *Proceedings of the North East Linguistics Society 24* (pp. 333-379). Amherst, MA: GLSA.
- Nazzi, T., Iakimova, G., Bertocini, J., Fredonie, S., & Alcantara, C. (2006). Early segmentation of fluent speech by infants acquiring French: Emerging evidence for crosslinguistic differences. *Journal of Memory and Language*, *54* (3), 283-299.
- Norris, D., McQueen, J. M., Cutler, A., & Butterfield, S. (1997). The possible-word constraint in the segmentation of continuous speech. *Cognitive Psychology*, *34*, 191-243.
- Pelucchi, B., Hay, J. F., Saffran, J. R. (2009). Statistical learning in a natural language by 8 month-old infants. *Child Development*, *80*(3), 674-685.
- Polka, L., & Sundara, M. (2003). Word segmentation in monolingual and bilingual infant learners of English and French. In M. J. Solé, D. Recasens, & J. Romero. (Eds.), *Proceedings of the 15th International Congress of Phonetic Sciences* (pp.1021-1024). Barcelona, Spain: Causal Productions.
- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*, 1926-1928.
- Seidl, A., & Johnson, E. K. (2008). Boundary alignment enables 11-month-olds to segment vowel initial words from speech. *Journal of Child Language*, *35*, 1-24.
- Shi, R., Cutler, A., Werker, J. F., & Cruickshank, M. (2006). Frequency and form as determinants of functor sensitivity in English-acquiring infants. *Journal of the Acoustical Society of America*, *119* (6), EL61-EL67.
- Shi, R., & Lepage, M. (2008). The effect of functional morphemes on word segmentation in preverbal infants. *Developmental Science*. *11* (3), 407-413.
- Shi, R., Werker, J., & Cutler, A. (2006). Recognition and representation of function words in English-learning infants. *Infancy*, *10* (2), 187-198.
- Thiessen, E. D., & Saffran, J. R. (2003). When Cues Collide: Use of Stress and Statistical Cues to Word Boundaries by 7- to 9-Month-Old Infants. *Developmental Psychology*, *39* (4), 706-716.
- Tranel, B. (1996). French Liaison and Elision Revisited: A Unified Account within Optimality Theory. In C. Parodi, C. Quicoli, M. Saltarelli & M. L. Zubizarreta (Eds.), *Aspects of Romance Linguistics* (pp. 433-455). Washington D.C.: Georgetown University Press.