INTERSPEECH 2019
September 15–19, 2019, Graz, Austria

Liquid deletion in French child-directed speech

Sharon Peperkamp, Monica Hegde, Maria Julia Carbajal

Laboratoire de Sciences Cognitives et Psycholinguistique, ENS - PSL University – EHESS – CNRS, Paris, France

sharon.peperkamp@ens.fr, hegdms0@gmail.com, carbajalmjulia@gmail.com

Abstract

In spoken language, words can have different surface realizations due to the application of language-specific phonological rules. Young children must acquire these rules in order to be able to undo their effects and recognize the intended words during language processing. Evidence so far suggests that they achieve this early on, but the learning mechanisms that they exploit are unknown. As a first step in examining this question, it is necessary to know to what extent phonological rules occur in their input. Here, we investigate the occurrence of liquid deletion, i.e. the optional deletion of the liquid in word-final obstruent-liquid clusters, in French child-directed speech. Analyzing a corpus from the Childes database that contains video recordings, we find that words finishing in obstruent-liquid clusters occur on average once every 13 utterances, and that in more than half of the cases the liquid is deleted. As in adult-directed speech, deletion applies more often before consonants than before vowels and pauses. Furthermore, pairs of tokens of the same word with and without deletion tend to cluster together, with a median distance of 49 seconds of speech. This clustering could be a powerful cue in the process of the acquisition of liquid deletion.

Index Terms: Child-directed speech, French, phonological rules, liquid deletion, phonological acquisition

1. Introduction

Language users have (mostly implicit) knowledge of the phonological rules of their native language, as evidenced by the phenomenon of compensation: during sentence processing, they undo the effects of phonological rules in order to retrieve the canonical form of words [1-5]. Young children thus must not only learn to apply the native phonological rules in their speech, they must also learn to compensate for them for the purposes of word recognition. Evidence so far suggests that they achieve this early on. Specifically, compensation for phonological rules during word recognition has been reported at the ages of 18 months (English flapping, [6]), 24 months (French voicing assimilation, [7]) and 33 months (French voicing and English place assimilation, [8]). Like in adults, this compensation is context-specific, as children do not undo the rule's effect when the triggering context is not present ([7,8]).

How phonological rules are acquired in perception is largely an open question. Most research so far has focused on analyzing the extent to which phonological rules apply in the child's input to begin with. It is well known that infant-directed speech has special properties that distinguish it from adult-directed speech, including shorter utterances and exaggerated prosody, properties that might enhance learnability (for a review, see [9]). Concerning phonological rules, one might expect that their rate of application is lower in child-directed than in adult-directed speech, because prosodic units are shorter and marked by stronger boundaries, which block the application of certain phonological rules, and/or because parents would favor careful speech and hence produce more canonical word pronunciations.

Several studies have examined phonological rules in child-directed speech and made a comparison with adult-directed speech. All of them involve consonant reduction and assimilation rules, mostly in English, with children's ages ranging from 3 months to almost 6 years. The results differ widely: while some studies indeed report lower application rates in child-directed than in adult-directed speech [10-13], others report higher [14] or comparable [15-18] application rates. The age of the children involved does not seem to explain these differences. The question of how the presence of a given rule might change as a function of the child's age, however, has been directly addressed in two studies. One of these was longitudinal, and found an increase followed by a decrease of a deletion rule in the input to toddlers between the ages of 1½ and 2½ years [19]; the other one was cross-sectional, and found an increase in the use of lenition rules between the ages of 2 and 4 years [11].

The above-mentioned studies also differ in the way the data were collected. For instance, recordings could be made in the lab or at home, and while the addressee in adult-directed speech could be a relative or a friend, in most studies it was an experimenter. Similarly, tasks for child-directed speech included picture- or storybook reading, playing with toys, and word teaching; those for adult-directed speech included reading, board-game playing, and interview. All of these aspects can influence the way parents speak, both to their child and to the interviewer. Lastly, while in most studies child- and adult-directed data were collected from the same adult participants, in some of them only child-directed speech was collected, with the comparison to adult-directed speech being made using existing corpus data [11,13,16].

In the present study we choose to focus on the type of child-directed speech that young children arguably hear the most. That is, we examine a CHILDES corpus of spontaneous interactions between parents and their child recorded at their home. Our case study is French liquid deletion (henceforth: LD), by which the liquid of word-final obstruent-liquid clusters is optionally deleted, e.g., *[tab]* < *[tabl]* 'table' or *[uv]* < *[uvʁ]* 'open$_{IMP}$'. While traditionally described as a rule that applies before consonants [20], two previous studies of adult-directed conversational speech found that LD also occurs before vowels and before pause, although less often than before consonants [21,22]. Thus, we consider all tokens of obstruent-liquid final words produced by an adult. Note that our choice of analyzing a CHILDES corpus implies that we cannot make a direct comparison between child- and adult-directed speech (the corpus does not contain enough adult-directed speech).

http://dx.doi.org/10.21437/Interspeech.2019-2838

Table 1: *Number of coded tokens and % liquid deletion split by context and speaker*

	father		mother		researcher		*TOTAL*
	nb tokens	**% LD**	**nb tokens**	**% LD**	**nb tokens**	**% LD**	***nb tokens***
C	315	79.1	296	81.4	52	84.6	*663*
pause	162	19.8	121	37.2	22	18.2	*305*
V	137	29.9	86	44.2	14	14.3	*237*
TOTAL	*614*	*53.4*	*503*	*65.7*	*88*	*55.3*	*1219*

However, we will compare our data to those of previous studies on adult-directed conversational speech [21,22] in the discussion.

Besides analyzing how many words ending in an obstruent-liquid cluster that are spoken by an adult occur in the corpus and how often these words undergo LD, we examine to what extent tokens of the same words with and without the final liquid co-occur in short stretches of speech. Indeed, it has been suggested that the clustering of canonical and rule-altered word tokens might attract infants' attention to the presence of a phonological rule [8,23]. In order for this to be the case, we should not only find that tokens with and without LD cluster together, but also that pairs of words that differ only in the presence or the absence of a final consonant (e.g., *forme* [fɔʁm] 'form' - *fort* [fɔʁ] 'strong') do not show the same amount of clustering.

2. Corpus study

2.1. Method

We analyzed the Yamaguchi corpus in the Childes database [24], which contains video recordings of spontaneous interactions between a single, male, child and his parents as well as a few other adults. We selected all recording sessions in which the child was aged between 1 and 4 years. Twenty-six recording sessions, in which the child ranged from age 1;03.9 to 3;11.13, were thus included.

All utterances spoken by an adult and containing a word ending in an obstruent-liquid cluster were identified in the orthographic transcriptions in the CHAT files. Using the online PhonBank interface, a coder listened to the audio of these utterances, and coded whether liquid deletion had occurred or not, and whether the utterance was directed to the child or to another adult. Whenever the coder was unsure, a second coder listened as well and in all cases a common decision could be made. The second coder also listened to the audio and coded the exact time from the beginning of the session at which each token occurred. The following additional variables were retrieved from the CHAT files: target word, phonological context (before consonant, (semi-)vowel, or pause[1]), and speaker identity.

[1] A target word was coded as prepausal only when it was utterance-final. Indeed, commas in the orthographic transcriptions were found not to correspond reliably to pauses (or even intonation breaks) in the audio. (There were no cases in which the target word preceded a colon or a semicolon.)

2.2. Results

The selected recording sessions contained 16,444 utterances produced by one of the adults, for a total of 1297[2] word tokens ending in an OL-cluster. Thus, on average a word-final OL-cluster occurred every 13 utterances. The distance between tokens (regardless of their lemma) across the corpus was, however, not normally distributed; the median was 4 utterances, with a maximum of 125.

Seventy-eight tokens could not be coded; 41 because of missing video, and 37 because of background noise, low volume, etc. Nearly all of the 1219 coded tokens were produced by the father, the mother, and the female researcher who was a friend of the family and present during the recording sessions. Fourteen tokens that were spoken by three different adults were excluded from analyses, leaving 1205 tokens for analyses. Among the included tokens, 89% were directed to the child. Given the low proportion of adult-directed tokens, the factor child- vs. adult-directed speech is not taken into account in any of the analyses below. However, all analyses were run a second time without the adult-directed tokens, yielding the same results.

Overall, 58.0% of the coded tokens had undergone liquid deletion. Table 1 shows both the number of coded tokens and the percentage of liquid deletion by context and by speaker.

The data contained 69 different lemmas, which showed a large amount of variation both in the number of coded tokens (ranging from 1 to 245) and in the percentage of LD (ranging from 0% to 96.3%). Figure 1 shows the number of tokens (top panel)[3] and the percentage of LD (bottom panel) for the 17 most frequent lemmas in the corpus; each of these lemmas has at least 16 coded tokens. The lack of a correlation between the number of tokens and the percentage of LD is striking.

[2] Seven additional utterances that, according to the written transcriptions contained a word with a final OL-cluster, were found to contain no such word.

[3] Two sessions contain a specifically high amount of the target word *mettre* 'to put'; one in which the family is packing moving boxes (22x) and one in which the child is unwilling to wear the microphone (33x).

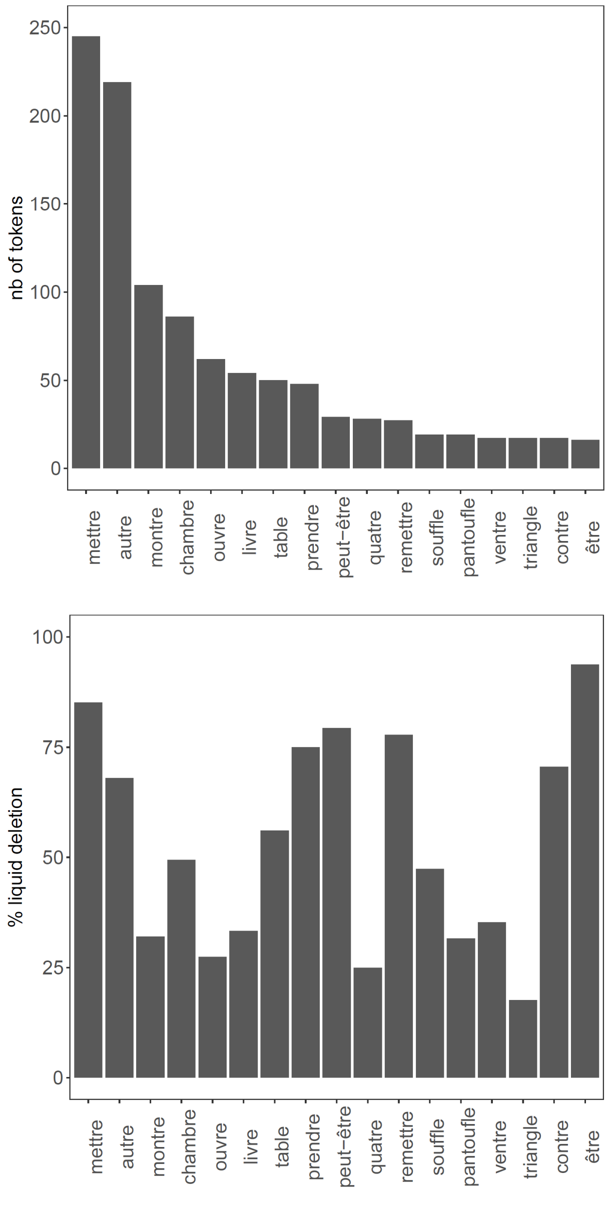

Figure 1: *Number of coded tokens (top) and percent LD (bottom) for the 17 most frequent lemmas in the corpus.*

The data were analyzed in a logistic mixed effects model using the *lme4* package [25] in R [26], with context, speaker, and their interaction as fixed factors (contrast-coded), and lemma as a random factor. Multiple comparisons were performed using the *lsmeans* package [27], which uses Tukey HSD correction. Both main effects and the interaction were significant: LD was more frequent before consonants than before either vowels ($\beta = 2.90$, SE = 0.35, $z = 8.35$, $p < .0001$) or pause ($\beta = 3.03$, SE = 0.30, $z = 8.35$, $p < .0001$), with no difference between the latter two contexts ($z < 1$), and the mother applied LD more often than either the father ($\beta = 0.42$, SE = 0.17, $z = 2.42$, $p < .05$) or the researcher ($\beta = 0.96$, SE = 0.39, $z = 2.46$, $p < .004$), with no difference between the latter two speakers ($\beta = 0.54$, SE = 0.39, $z = 1.39$, $p > .01$). The interaction was due to the fact that the mother's higher application rate was manifest before vowels when compared to the researcher and before pause when compared to the father.

We also examined whether the rate of LD changed over time. Figure 2 shows the percentage of LD per session.

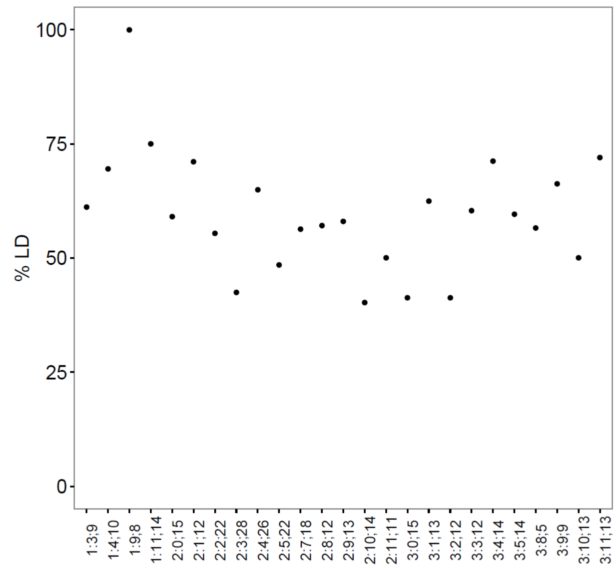

Figure 2: *Percent LD over time.*

There was no correlation between the child's age and the percent LD (adjusted $R^2 = 0.04$, $F(1, 23) = 1.84$, $p > .1$). The peak in the third session (age 1;9:8) is due to the fact that on PhonBank only the first 3 minutes of this session contain transcriptions, yielding a single datapoint.

Finally, we analyzed the distribution of distance between forms with and without the final liquid of each lemma, in order to obtain a measure of their clustering in time. Within each file, we searched for all consecutive pairs of the same lemma (not necessarily occurring within the same sentence), as shown in Figure 3. Then, for each pair, we coded the surface forms of the tokens as *same* (i.e., both with or without the liquid) or *different* (i.e., one with and one without the liquid). For those pairs coded as *different*, we computed their distance in seconds as well as in number of utterances.

Figure 3: *Example of consecutive tokens with same lemma. White rectangles indicate that the token was produced with a final liquid, while grey rectangles indicate that the liquid was omitted. A dashed connecting line indicates a pair coded as <u>same</u>, while a full line indicates a pair coded as <u>different</u>.*

Throughout the whole corpus, we found 920 token pairs. Of these, 57 (6%) could not be coded due to missing information regarding the surface form of one or both tokens, and were thus excluded from the analysis. Of the remaining 863 pairs, 318 (37%) were coded as *different*, and 545 (63%) as *same*. The distribution of distances for *different* pairs had a median of 49s (min = 0s, max = 3375s), or, equivalently, 12 utterances (min = 0, max = 589). Thus, almost 40% of all consecutive token pairs with the same lemma have different surface forms, and of these, over half occur within 1 minute of speech.

In order to evaluate whether the clustering of LD tokens with *different* surface forms could be a relevant cue to learn the phonological rule, we compared their distribution to that of true phonological neighbors differing only in the presence or absence of a final consonant. Within each file, we searched for all words finishing in a C_1C_2 cluster (e.g., *forme* [fɔʁm] 'form'), and their respective phonological neighbors differing only in the absence of C_2 (in this example, *fort* [fɔʁ] 'strong'). As these pairs were not time-coded, we used number of utterances between both tokens as the distance measure. Throughout the whole corpus, we found a total of 248 pairs of consecutive phonological neighbors. The distribution of distances for these pairs had a median of 39 utterances (min = 0, max = 497), that is, more than three times the median utterance distance between LD pairs with same lemma but different form. The difference in their distributions of utterance distances was significant (two-tailed Wilcoxon-Mann-Whitney test, $W = 28810$, $p < 0.0001$), indicating a higher amount of clustering for tokens differing in surface form due to LD than for true phonological neighbors.

3. Discussion

Analyzing the Yamaguchi corpus of French child-directed speech [24], we found that LD occurs frequently, with an average rate of 58%. LD is more frequent in the mother's speech than in that of the father and the experimenter, but all three speakers apply LD much more often before consonants than before vowels and pause, as has also been reported for adult-directed speech [21,22]. As to the potential differences between child- and adult-directed speech, we could not analyze this issue directly, due to the relatively small amount of adult-directed utterances in the corpus. However, we can compare our data to the ones obtained in the two previous corpus studies of adult-directed speech [21,22]. Both considered regional, age and gender variation while the first one also compared read to conversational speech. In this first study, both region and speech style showed effects. We therefore analyzed the subset of their data from conversational speech by speakers in France (as opposed to Belgium and Switzerland), and found an application rate of 51%. The second study examined LD in conversational speech only. No statistical analyses were carried out, but the data suggests differences in gender and socioeconomic status. For middle-class speakers, who match the socioeconomic status of the parents in the Yamaguchi corpus (N. Yamaguchi, personal communication), the application rate was 46%.

Thus, both studies found numerically lower rates of LD than the one we observed. It is hard to conclude from these comparisons that LD occurs more frequently in child- than in adult-directed speech. Indeed, the child-directed speech consisted of spontaneous interactions at home, with the researcher being a friend of the family, while the adult-directed speech had been collected in semi-directed sociolinguistic interviews; this might account for the relative higher LD rate in the child-directed data. If anything, the fact that our analyses did not differ when we omitted the adult-directed utterances leads us to the tentative conclusion that child- and adult-directed speech do not differ in the amount of LD.

As to the potential difference between male and female speech, no firm conclusion can be drawn either. Recall that in one of the studies on adult-directed speech, there was no effect of gender [22], while the other one suggested more LD in male than in female speakers [21]. Our own analyses of child-directed speech showed the highest LD rate in one of the female speakers. Based on our impressionistic observations we speculate that this speaker had a considerably higher speech rate, which might account for this finding. Clearly, data collection with many more speakers is necessary to determine whether male and female speakers differ in their amount of LD application, and whether such a difference is modulated by the type of speech, i.e. adult- vs child-directed.

Finally, we found that canonical and rule-altered tokens of the same word typically co-occur in short stretches of speech, whereas this is not the case for pairs of different words that also differ only in the presence or absence of a final consonant, such as *forme* [fɔʁm] 'form' – *fort* [fɔʁ] 'strong'. This might attract young children's attention to the presence of LD and help them acquire the rule. It should be noted that two alternative mechanisms for the acquisition of phonological rules are implausible. The first of these is based on word meaning: after having learned the meaning of at least a few words that undergo the rule, infants might note that these words have two phonological forms. As some knowledge about phonological rules has been shown to be present already at the age of 14 months [23], we deem this scenario in general unlikely. Alternatively, it has been hypothesized that before having access to word meanings but once they can segment words out of continuous speech, infants might exploit distributional information about word *forms*, by observing that certain pairs of minimally different word forms occur in different phonological contexts [28]. While this scenario has not been tested, there is evidence from artificial language-learning that twelve-month-old infants have the prerequisite, i.e. they can track the distribution of minimally different word forms [29]. In the case of LD, though, which occurs in all contexts, a distributional analysis simply fails to work. Thus, it is only the clustering properties we discovered that provide a robust distributional cue to the distinction between rule-based pairs of word forms (such as *tricycle – tricyc')* and similar pairs of minimally different word forms that are not linked by a rule (such as *forme – fort*). This, then, allows infants to potentially acquire LD before they know any words.

To conclude, we have shown that LD is pervasive in spontaneous child-directed speech in the home environment, and that the clustering of tokens with and without deletion of the same words could allow infants to acquire the rule without having to rely on word meanings. Future research should test if infants are sensitive to this clustering, and at what age they acquire LD and start compensating for it during sentence processing, like adults do [30].

4. Acknowledgements

This study was supported by ANR-17-EURE-0017 and ANR-17-CE28-0007-01. We thank Naomi Yamaguchi for practical assistance, and Mathieu Avanzi for sharing the raw data of [22] with us.

5. References

[1] M.G. Gaskell and W.D Marslen-Wilson, "Phonological variation and inference in lexical access." *Journal of Experimental Psychology: Human Perception and Performance*, vol. 22, pp. 144-158, 1996.

[2] E. Coenen, P. Zwitserlood, and J. Bölte, "Variation and assimilation in German: Consequences for lexical access and representation." *Language and Cognitive Processes*, vol. 16, no. 5-6, pp. 535-564, 2001.

[3] M.G. Gaskell and N.D. Snoeren, "The impact of strong assimilation on the perception of connected speech." *Journal of Experimental Psychology: Human Perception and Performance*, vol. 34, no. 6, pp. 1632-1647, 2008.

[4] N.D. Snoeren, J. Segui, and P. Hallé, "Perceptual processing of partially and fully assimilated words in French." *Journal of Experimental Psychology: Human Perception and Performance*, vol. 34, no. 1, pp. 193-204, 2008.

[5] H. Mitterer and J. McQueen, "Processing reduced word-forms in speech perception using probabilistic knowledge about speech production." *Journal of Experimental Psychology: Human Perception and Performance*, vol. 35, no. 1, pp. 244–263, 2009.

[6] A. Chong and M. Sundara, "18 Month-olds compensate for a phonological alternation," in E. Grillo and K. Jepson (eds.), *Proceedings of the* 39th annual *Boston University Conference on Language Development*, pp. 113-126, 2015.

[7] K. Skoruppa, N. Mani, K. Plunkett, D. Cabrol and S. Peperkamp, "Early word recognition in sentence context: French and English 2-year-olds' sensitivity to sentence-medial mispronunciations and assimilations," *Infancy*, vol 18, 1007-1029, 2013.

[8] K. Skoruppa, N. Mani, and S. Peperkamp, "Toddlers' processing of phonological alternations: Early compensation for assimilation in English and French," *Child Development*, vol 84, pp. 313-330, 2013.

[9] M. Soderstrom, "Beyond babytalk: Re-evaluating the nature and content of speech input to preverbal infants," *Developmental Review*, vol 27, no. 4, pp. 501–532, 2007.

[10] N. Bernstein Ratner, "Phonological rule usage in mother–child speech," *Journal of Phonetics*, vol 12, 245–54, 1984.

[11] P. Foulkes, G.J. Docherty, and D. Watt, "Phonological variation in child-directed speech," *Language*, vol. 81, no. 1, pp .177-206, 2005.

[12] L.C. Dilley, A.L Millett, J.D. McAuley, and T.R. Bergeson, "Phonetic variation in consonants in infant-directed and adult-directed speech: the case of regressive place assimilation in word-final alveolar stops," *Journal of Child Language*, vol. 41, no. 1, pp. 155-175, 2014.

[13] H. Andreassen, "Schwa. Distribution and acquisition in light of Swiss French data," Ph.D. dissertation, University of Tromsø, 2013.

[14] L. Shockey and Z.S. Bond, "Phonological processes in speech addressed to children," *Phonetica*, vol. 37, no. 4, pp. 267-274. 1980.

[15] L. Fais, S. Kajikawa, S. Amano, and J.F. Werker, "Now you hear it, now you don't: vowel devoicing in Japanese infant-directed speech," *Journal of Child Language*, vol 37, no. 2, pp. 319-340, 2010.

[16] K. Miller. "Acquisition of variable rules:/s/-lenition in the speech of Chilean Spanish-speaking children and their caregivers," *Language Variation and Change*, vol 25, no. 3, pp. 311-340, 2013.

[17] M. Lahey and M. Ernestus, "Pronunciation variation in infant-directed speech: Phonetic reduction of two highly frequent words," *Language Learning and Development*, vol 10, 308-327, 2014.

[18] H. Buckler, H. Goy, and E.K. Johnson. "What infant-directed speech tells us about the development of compensation for assimilation," *Journal of Phonetics*, vol 66, pp. 45-62, 2018.

[19] H. Buchan and C. Jones, "Phonological reduction in maternal speech in northern Australian English: change over time." *Journal of Child Language*, vol 41, no. 4, pp 725-755, 2014.

[20] F. Dell, "Consonant clusters and phonological syllables in French." *Lingua*, vol 95, pp. 5-26. 1985.

[21] Z. Boughton, "Social class, cluster simplification and following context. Sociolinguistic variation in word-final post-obstruent liquid deletion in French," *Journal of French Language Studies*, vol 25, pp 1-21, 2015.

[22] S. Brognaux and M. Avanzi, "Sociophonetics of phonotactic phenomena in French," *ICPhS 2015 - 18th International Congress of Phonetic Sciences, August 10-14, Glasgow, UK, Proceedings*, pp. 0744. 2015.

[23] P. Fort, J. Brusini, Y Carbajal, Sun, and S. Peperkamp, "A novel form of perceptual attunement: context-dependent perception of a native contrast in 14-month-old infants," *Developmental Cognitive Neuroscience*, vol. 26, pp. 46-51, 2017.

[24] N. Yamaguchi, "Parcours d'acquisition des sons du langage chez deux enfants francophones," Ph.D. Dissertation, Université Sorbonne Nouvelle Paris 3, 2012.

[25] D. Bates, M. Maechler, B. Bolker & S. Walker, "Fitting linear mixed-effects models using lme4," *Journal of Statistical Software*, vol. 67, 1-48, 2015.

[26] R Core Team, "R: A language and environment for statistical computing," R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/, 2014.

[27] R. Lenth, "Least-squares means: The R package lsmeans," *Journal of Statistical Software*, vol. 69, pp 1-33, 2016.

[28] S. Peperkamp and E. Dupoux, "A typological study of stress 'deafness'," in C. Gussenhoven and N. Warner (eds.), *Laboratory Phonology VII*, Berlin: Mouton de Gruyter, 203-240, 2002.

[29] K. White, S. Peperkamp, C. Kirk, and J. Morgan, "Rapid acquisition of phonological alternations by infants," *Cognition*, vol 107, 238-265, 2008.

[30] S. Peperkamp and A. Martin Iturralde Zurita, "Compensation for French liquid deletion during auditory sentence processing", *Proceedings of Interspeech*, 2019.