COGNITION

# The labial–coronal effect revisited: Japanese adults say pata, but hear tapa

Sho Tsuji [a],[*],[1], Nayeli Gonzalez Gomez [b],[2], Victoria Medina [a],[1], Thierry Nazzi [b],[c],[2], Reiko Mazuka [a],[d],[1],[3]

[a] Laboratory for Language Development, RIKEN Brain Sciences Institute, 2-1 Hirosawa, Wako City, Saitama 351-0198, Japan
[b] Université Paris Descartes, Sorbonne Paris Cité, 45, rue des Saints-Pères, 75006 Paris, France
[c] CNRS, Laboratoire Psychologie de la Perception, Paris, France
[d] Duke University, 417 Chapel Drive, Box 90086, Durham, NC 27708, USA

## ARTICLE INFO

## ABSTRACT

The labial–coronal effect has originally been described as a bias to initiate a word with a labial consonant–vowel–coronal consonant (LC) sequence. This bias has been explained with constraints on the human speech production system, and its perceptual correlates have motivated the suggestion of a perception–production link. However, previous studies exclusively considered languages in which LC sequences are globally more frequent than their counterpart. The current study examined the LC bias in speakers of Japanese, a language that has been claimed to possess more CL than LC sequences. We first conducted an analysis of Japanese corpora that qualified this claim, and identified a subgroup of consonants (plosives) exhibiting a CL bias. Second, focusing on this subgroup of consonants, we found diverging results for production and perception such that Japanese speakers exhibited an articulatory LC bias, but a perceptual CL bias. The CL perceptual bias, however, was modulated by language of presentation, and was only present for stimuli recorded by a Japanese, but not a French, speaker. A further experiment with native speakers of French showed the opposite effect, with an LC bias for French stimuli only. Overall, we find support for a universal, articulatory motivated LC bias in production, supporting a motor explanation of the LC effect, while perceptual biases are influenced by distributional frequencies of the native language.

© 2012 Elsevier B.V. All rights reserved.

## 1. Introduction

Some speech sounds and speech sound patterns are more frequent than others across languages. For instance, all languages archived by linguists possess plosives like /t/ and /d/ and syllables with a consonant–vowel (CV) structure, while not all languages possess plosives /p/ or syllables with a CVC structure (cf. Locke, 2000; Maddieson, 1984). Such cross-language-commonalities have been attributed to biological restrictions on language production and perception on phylogenetic and ontogenetic scales (Locke, 2000; MacNeilage & Davis, 2000).

The labial-coronal (LC) bias describes a predominance of labial–coronal consonant sequences (e.g., /pata/) compared to coronal–labial ones (CL, e.g., /tapa/) in CVC or CVCV sequences (MacNeilage & Davis, 2000; MacNeilage, Davis, Kinney, & Matyear, 1999). This bias has been found in many languages, although it has been suggested that Japanese and Swahili might be exceptions (MacNeilage et al., 1999). It has also been found to influence infants' early words (MacNeilage et al., 1999), and both adult speech production (Rochet-Capellan & Schwartz, 2007) and adult speech per-

* Corresponding author. Present address: Max-Planck Institute for Psycholinguistics, PO Box 310, 6500 AH, Nijmegen, The Netherlands. Tel.: +31 24 3612160; fax: +31 24 3521213.

E-mail addresses: Sho.Tsuji@mpi.nl (S. Tsuji), naygon@hotmail.com (N.G. Gomez), medina_vicky@yahoo.fr (V. Medina), thierry.nazzi@paris-descartes.fr (T. Nazzi), mazuka@brain.riken.jp (R. Mazuka), mazuka@brain.riken.jp (R. Mazuka).
[1] Tel./fax: +81 48 467 9770.
[2] Tel.: +33 1 42 86 43 15; fax: +33 1 42 86 33 22.
[3] Tel.: +1 919 660 5702; fax: +1 919 660 5726.

ception (Sato, Vallee, Schwartz, & Rousset, 2007). Given the pervasiveness of this bias (which however was mostly studied in plosive consonants), it was proposed to result from motor constraints of the human production system: the relative ease at producing LC sequences compared to CL sequences would translate into a higher frequency of LC sequences in the lexicon of most languages, and biases in both perception and production of these sequences.

In this context, and in spite of the dominating tendency for an LC bias in the languages investigated, Japanese has been pointed out as an exception to this pattern: MacNeilage et al. (1999) claimed that in Japanese, CL sequences occur more frequently than LC sequences (MacNeilage et al., 1999). If this were true, this would suggest that motor constraints behind the LC bias could be modulated or even reversed in the lexicon of a language, which would then raise issues regarding how speakers of that language perceive and produce LC and CL sequences. This finding was, however, based on a very small sample of words. Therefore, the present research will first examine the distribution of LC and CL sequences in the adult Japanese lexicon based on two large samples of Japanese discourses (Corpus Analysis). These analyses will bring detailed information regarding the "exceptional" status of Japanese in terms of the LC bias. This will allow us, second, to explore if and how perception and production of LC and CL sequences are biased in adult speakers and listeners of this language, and use these data to evaluate the motor and perceptual explanations previously offered to explain the LC bias. Before presenting the results of our research, the remainder of the introduction will summarize previous research on the production and perception of LC and CL sequences.

### 1.1. The LC bias in production

The LC bias was initially reported in young children's early productions. Ingram (1974) reported one English- and one French-learning infant's tendency to initiate words with a labial consonant, followed by a consonant in posterior position. Similarly, Locke (2000) reported an "anterior-to-posterior progression" in young children's productions. Looking at a larger sample, MacNeilage et al. (1999) analyzed plosive /p, b, t, d/ and nasal /m, n/ segments in the first words of 10 English-learning infants, finding an LC bias in nine of them and an overall ratio of LC–CL sequences of 2.55. The prevalence of this bias across languages was confirmed in a review of seven studies focusing on infants' early productions in English, German, Dutch, French, and Czech (MacNeilage & Davis, 1998). A longitudinal analysis of five Dutch-learning children suggests that the early LC bias is associated with a certain developmental stage: Fikkert and Levelt (2008) report that Dutch children, as soon as they start combining consonants with different place of articulation features in production, go through a stage in which they preferably produce LC sequences.

This LC bias is also reflected in the inventories of languages. Lexicon counts of ten languages (English, Estonian, French, German, Hebrew, Japanese, Maori, Quichua, Spanish, and Swahili) revealed an overall ratio of LC–CL sequences of 2.23 (MacNeilage et al., 1999). Except for Swahili and Japanese, the lexicon counts of all languages

clearly revealed a higher frequency of LC compared to CL sequences, with only Japanese showing a trend in the opposite direction. However, the results obtained for some languages were based on very small samples of words. In particular, the Japanese data were based on 68 words extracted from a travel dictionary, which makes it necessary to reassess these results.

Several motor accounts have been proposed for the observed LC effect in language inventories as well as in language learning. The first one is based on the, possibly self-organizational, tendency of infants to start out an utterance with an easy element and then add complexity (MacNeilage & Davis, 2000). In the *frame-content* theory, a labial CV sequence is defined as the default, *pure frame* resulting from simple mandibular oscillation, while a coronal CV sequence or *fronted frame* requires an additional tongue movement. Alternatively, Rochet-Capellan and Schwartz (2007) proposed that LC sequences have a higher articulatory stability than CL sequences. Their criticism of the 'simple first' account includes that it is not clear if labial sequences are easier to produce than coronal ones (Vilain, Abry, Badin, & Brosda, 1999), and that a developmental explanation is not sufficient to explain the persistence of the LC bias in adult lexicons. In order to assess articulatory stabilities, French participants were asked to repeat LC and CL sequences (/pata/ and /tapa/, /pasa/ and /sapa/, /fata/ and /tafa/) in a speeded articulation task. The results first showed that speeding leads to a shift from one jaw cycle per syllable to one per disyllable through vowel reduction after one of the consonants, so that an initial CVCV sequence evolves into a CCV cluster (e.g., /pata/→/p'ta/). Second, shifts to an LC sequence like /p'ta/ were favored over shifts to a CL sequence like /t'pa/ for speeded LC and CL sequences (e.g., /pata/→/p'ta/; /tapa/→/p'ta/), suggesting a higher coordinative stability for LC compared to CL sequences.

While the above two accounts differ widely in the processes they suggest as the cause of the LC bias, they share the assumption that it is located in properties of the human speech production system. For this common assumption, a test of the sequence preferences in speakers of a language with different distributions would be crucial.

### 1.2. The LC bias in perception

Previous findings suggest that the articulatory stability of speech forms is coupled to their perceptual stability (Sato, Schwartz, Abry, Cathiard, & Loevenbruck, 2006). For example, the articulatory more stable CCV sequence /ps/ +vowel shows a higher perceptual stability than the less stable CVC sequence /s/+vowel+/p/. These findings motivated the study of possible perceptual correlates of the LC bias (Sato et al., 2007). To this end, the verbal transformation effect, a multistability perception phenomenon describing changes in perception during listening to the continuous rapid alternation of a speech form (Warren, 1961; Warren & Gregory, 1958), was exploited. For instance, while listening to rapid repetitions of the word "rest", listeners are likely to switch between perceiving it as a repetition of "rest" and "tress" or "stress" (Warren & Gregory, 1958).

French adults were presented with rapid repetitions of LC and CL sequences in voiceless (/p/,/t/), or voiced (/b/,/d/) plosive consonant contexts, and in the vowel contexts /a/, /i/, or /o/. Importantly, a lexical analysis showed an LC bias of 1.43 for voiceless or voiced plosive sequences taken together and for the subset of voiceless plosives (2.32), but a CL bias for voiced plosives (0.57), so that, from an input perspective, diverging perceptual biases were a possible outcome for these subsets. Hence, rather than being a consequence of motor constraints, the tendency to perceive LC rather than CL in the verbal transformation task might be a direct result of the input.

The ratio of time participants spent perceiving the sequences as LC or CL was calculated as an index of perceptual stability. Results showed that LC sequences were more stable than CL sequences for both voiceless and voiced plosives. Therefore, the results did not reflect the reversed input CL bias of voiced plosives. French listeners appeared to have an overall LC preference, at least for plosives. The authors interpret the results in the context of a perception–action link (e.g., Liberman & Whalen, 2000; Schwartz, Basirat, Ménard, & Sato, 2012), suggesting that the articulatory advantage of LC chunking is connected to its perceptual chunking.

However, an influence of input on the LC bias as an alternative explanation cannot be discarded: The LC bias in French is true both overall but also restricted to sequences of all plosives in French (Sato et al., 2007; Vallée, Rousset, & Boë, 2001), and this strong bias could override the local CL bias restricted to voiced plosives (see also infant data below). This would be in line with numerous studies showing that ambient language structures affect segmentation, both in studies of natural language segmentation (e.g., McQueen, 1998; Weber & Cutler, 2006), or artificial language segmentation (e.g., Mersad & Nazzi, 2011; Pena, Bonatti, Nespor, & Mehler, 2002; Saffran, Newport, & Aslin, 1996). These ambient language influences can also be observed in infants, who start preferring to listen to words with legal over illegal phonotactic patterns in their native language (e.g., Friederici & Wessels, 1993; Jusczyk, Friederici, Wessels, Svenkerud, & Jusczyk, 1993), and frequent over infrequent speech sound sequences (Jusczyk, Luce, & Charles-Luce, 1994) between 6 and 9 months of age.

Taking this input-based alternative into account, Nazzi, Bertoncini, and Bijeljac-Babic (2009) tested the LC bias in French-learning infants of 6 and 10 months of age. In a head-turn preference paradigm (HPP), infants were tested on their preference for lists of LC vs. CL CVCV sequences that included both voiceless and voiced plosives, showing they preferred the LC lists at 10, but not 6, months. These results strongly suggest that language input might play a role in infants' development of a perceptual LC bias. This finding was extended to CVC plosive sequences, showing the emergence of an LC bias between 7 and 10 months of age (Gonzalez-Gomez & Nazzi, 2012b). In both infant studies, as in Sato et al. (2007), plosive consonants were used. In order to extend these findings to other manners of articulation and to further explore the level on which input biases influence perception, Gonzalez-Gomez and Nazzi (2012a) later explored the presence of the LC bias in the adult lexicon at a more fine-grained level. After establishing that the overall

LC bias is found on sequences restricted to two manners of articulation (sequences of plosives and sequences of nasals) but not to sequences of fricatives, they tested French-learning 10-month-olds on LC versus CL preferences for the three different manners of articulation separately. The results showed an LC bias for plosives and nasals, and the opposite CL bias for fricatives, lending further support to an input-based origin of the LC bias in perception (but see Rochet-Capellan & Schwartz, 2007, for a discussion of motor specificities that could lead to differences between plosives and fricatives) that is learned at the level of classes of consonants defined by manner of articulation.

The above findings underline the importance of further exploring the LC bias, at different ages (infants, adults), in different languages (that have an overall LC bias, as all languages studied so far, or that have been proposed to have an overall CL bias), and possibly also for different classes of consonants. In particular, in order to tease apart the motor and perceptual explanations, it is important to test the LC bias in cases in which the adult input has a CL bias either overall or in the subset of plosives, since this is the manner that has been discussed most extensively in the context of an LC bias (cf. MacNeilage et al., 1999; Sato et al., 2007). In such a case, motor explanations still predict an LC bias while perceptual explanations predict a CL bias as was observed, for example, for fricatives in French infants (Gonzalez-Gomez & Nazzi, in press). The present study was intended to start testing such cases in Japanese adults.

### 1.3. Aims of the current study

The current study assesses articulatory and perceptual biases in adult speakers of Japanese. As a first step, the trend towards a dominance of CL over LC sequences in the adult lexicon (MacNeilage et al., 1999) was reassessed by analyses of large corpora of Japanese (see Section 2. Corpus Analysis). Given the results by Gonzalez-Gomez and Nazzi (2012a), these analyses were conducted either overall, or separated by manner of articulation. Based on our findings, we subsequently studied Japanese adults' articulatory and perceptual preferences for LC versus CL sequences for a subset of consonants that exhibits a CL bias in the adult lexicon.

In order to compare our results to previous studies, the design of the production study (Experiment 1) was closely matched to Rochet-Capellan and Schwartz (2007). The perceptual experiment was also closely matched to Sato et al. (2007). However, we extended it by using a fully crossed design (presenting Japanese adults with stimuli recorded by a Japanese speaker and stimuli recorded by a French speaker in Experiment 2, and then presenting French adults with the same stimuli in Experiment 3) in order to explore possible effects of language of presentation, and to replicate the previous results in French with our new set of stimuli. Lastly, Experiment 4 addressed some possible interpretations of the language-of-presentation effects found in Experiments 2 and 3.

If the LC biases found in production and perception are due to articulatory factors, then Japanese participants are expected to show a higher articulatory stability of LC compared to CL sequences, and both Japanese and French adults should show an LC preference in perception. However, if

preferences are influenced by input properties, because the consonants tested have a CL input bias in Japanese but an LC input bias in French, Japanese participants are expected to show a CL bias both in production and perception, while French participants should have a perceptual LC bias. Note that while the above predictions are made for both production and perception, it remains possible that dissociations will be observable, the present study being the first one to try to directly articulate the link between input properties, production and perception biases in determining the LC bias.

## 2. Corpus analysis

In order to reassess the findings of MacNeilage and Davis (2000) that Japanese has a higher frequency of CL compared to LC patterns, two large adult corpora of Japanese were analyzed.

Given the manner of articulation effects found in the developmental studies by Gonzalez-Gomez and Nazzi (2012a), we conducted a series of analyses with all manners together, and two other series of analyses restricted to either plosives or nasals. The reason for not exploring distributions for the other manners of articulation independently was dictated by the fact that in Japanese labial fricatives are extremely rare. Importantly though, the LC effect has originally and predominantly been assessed with plosives and nasals (i.e., MacNeilage, Davis, Kinney, and Matyear, 2000; Sato et al., 2007), as these are the first sounds produced by infants. Therefore, looking at the patterns for these subgroups separately is especially important.

### 2.1. Input corpora

As a corpus of written language, the NTT frequency corpus (Amano & Kondo, 2000) was chosen, which contains all written content of the Asahi Newspaper, a major Japanese daily newspaper, over 14 years (1985–1998). The original written text includes the three Japanese script types *kanji, hiragana* and *katakana*, as well as some alphabetic scripts. Katakana transcriptions for all forms except the alphabetic scripts are provided, which allowed us to do an unambiguous phonemic transcription of the segments of interest.

As a corpus of spoken language, the subsection 'simulated public speech' of the Corpus of Spontaneous Japanese (CSJ) (Maekawa, 2003) was chosen. It includes speech of 590 participants holding a 10–12-min speech on an everyday topic in front of a small audience. The corpus used for the analyses, includes phonemic transcriptions by trained phoneticians.

The target consonants for the analyses of all manners were labial (p,b,m,f,v) and coronal (t,d,n,s,z,ʃ,tʃ,j,r) segments. Note that the labial segments (f,v) are very low-frequency segments (with the exception of /f/ in front of the vowel /u/, they appear exclusively in recent loanwords). For the analyses of plosives, we used labial (p, b) and coronal (t, d) plosives, and for the analyses of nasals, we used labial (m) and coronal (n) nasals. All CVC sequences were analyzed to determine the token frequencies of LC and CL

sequences. Three different frequency analyses were conducted: Firstly, for any CVC sequence within a word (ANY); secondly, for word-initial CVC sequences only (INI); and thirdly, for CVCV words only (WORD).

### 2.2. Results and discussion

Results of the analyses are presented in Table 1. Chi-square tests were conducted to test for the significance of the differences between LC and CL occurrences.

The first remarkable finding is that overall, very similar results are obtained for the two corpora, which suggests that the effects found are robust. Indeed, the few differences observed are due to differences in the size of the biases, while the direction of the biases observed is always the same across the two corpora. Second, it also appears that the results are not affected by the positions/structures of sequences we analyzed, since similar results are found whether the analyses are performed anywhere within a word (ANY), word-initially (INI), or in words with a CVCV structure (WORD). This suggests that the constraints that apply to labial and coronal sequences are very strong and independent of their position with respect to word boundaries.

Regarding the bias itself, it is noteworthy that our findings do not support the claim by MacNeilage et al. (1999), based on a very small sample of 68 Japanese words, that Japanese is a language with a CL bias. On the contrary, it appears that, overall, Japanese is a language with an LC bias, like most other languages reported so far.

However, the overall bias translates differently for the two manners of articulation on which restricted analyses could be conducted. For nasals, the LC–CL ratios were above 1 for both corpora and for the ANY, INI and WORD analyses, with significant differences between frequencies of LC and CL occurrences. But for plosives, the LC to CL ratios were below 1 for all 6 comparisons, indicating a higher frequency of CL compared to LC sequences. Chi-square tests indicate that the difference between LC and CL frequencies are statistically significant for all comparisons.

In summary, the adult Japanese lexicon thus has an overall LC bias, while a CL bias was found but only restricted to sequences of plosives. On the one hand, these results support the notion of a universal LC bias, and Japanese is no exception to this pattern. On the other hand, Japanese deviates from this overall pattern with regard to plosive sequences.[4] As this is the manner that has been focused on in many previous studies on the LC bias (MacNeilage et al., 1999; Sato et al., 2007), Japanese is, despite its lack of an overall LC bias, an ideal test case for the current research, because it shows a CL bias in this most critical manner. Given these findings, and since we were interested in determining Japanese adults' articulatory and perceptual preferences in cases in which there was a CL bias in the input, for which motor and perceptual explanations of the LC bias make different predictions, the remainder of our study focused on compar-

---

[4] While in French, the overall LC bias was found also on the analyses restricted to plosives and nasals, but was reversed for fricatives (Gonzalez-Gomez & Nazzi, 2012a).

**Table 1**
Token frequencies, LC/CL ratios and chi-square tests of plosive, nasal and all LC and CL sequences in the NTT and CSJ corpora.

| | | ANY | | INI | | WORD | |
|---|---|---|---|---|---|---|---|
| | | NTT | CSJ | NTT | CSJ | NTT | CSJ |
| All manners | LC | 20,762,465 | 211,897 | 13,276,873 | 156,841 | 6,298,998 | 105,756 |
| | CL | 11,209,479 | 84,867 | 5,594,640 | 44,662 | 3,086,723 | 12,199 |
| | Ratio | 1.85 | 2.50 | 2.37 | 3.51 | 2.04 | 8.67 |
| | $\chi^2$ | 123780844.4 | 1119548.5 | 131252.8 | 2624008.5 | 293129772.5 | 2732623.2 |
| | $p$ | <.001 | <.001 | <.001 | <.001 | <.001 | <.001 |
| Plosives | LC | 437,106 | 3015 | 137,607 | 1360 | 9,627 | 94 |
| | CL | 567,420 | 7682 | 236,449 | 5264 | 103,975 | 2156 |
| | Ratio | 0.77 | 0.39 | 0.58 | 0.26 | 0.09 | 0.04 |
| | $\chi^2$ | 202697.0 | 4331.8 | 134382.6 | 5745.6 | 4820689.6 | 1978.7 |
| | $p$ | <.001 | <.001 | <.001 | <.001 | <.001 | <.001 |
| Nasals | LC | 7681 | 28,038 | 2,702,830 | 23,837 | 802,292 | 12,480 |
| | CL | 4315 | 4572 | 328,112 | 2879 | 152,237 | 971 |
| | Ratio | 1.78 | 6.13 | 8.24 | 8.28 | 5.27 | 12.85 |
| | $\chi^2$ | 2625.7 | 120440.3 | 17187075.1 | 152566.1 | 2775747.7 | 136413.2 |
| | $p$ | <.001 | <.001 | <.001 | <.001 | <.001 | <.001 |

ing Japanese adults' production and perception of LC and CL sequences restricted to plosive consonants.

## 3. Experiment 1: Production

This experiment assesses the relative articulatory stability of plosive LC versus CL disyllables in speakers of Japanese. A previous study in French found that the speeded production of LC and CL plosive CVCV sequences evolves more frequently towards CCV sequences with an LC consonant cluster than towards one with a CL cluster (Rochet-Capellan & Schwartz, 2007). Since for plosives, and contrary to French, we found that CL sequences are more frequent than LC sequences in Japanese, it was of interest if a similar LC articulatory pattern would be found for Japanese adults (motor interpretation), or whether they would show a CL bias (perceptual interpretation). Procedure and analysis were closely matched to Rochet-Capellan and Schwartz (2007).

### 3.1. Participants

Nineteen undergraduate students (seven females) of a Japanese university in the Tokyo area with a mean age of 19.7 years (range: 19–22) participated in the experiment for payment. All speakers were native speakers of Japanese without speech or hearing problems.

### 3.2. Stimuli

The phonetic material to be produced consisted of four LC (/pata/, /pete/, /piti/, and /putu/), and four CL (/tapa/, /tepe/, /tipi/, and /tupu/) CVCV disyllables. While Rochet-Capellan and Schwartz (2007) employed plosive and mixed plosive-fricative consonants in the vowel context /a/, we restricted our stimuli to the plosive manner of articulation for two reasons: because a CL advantage was only found restricted to plosives in Japanese, and because there are no labial fricatives in Japanese. Given this, and in order to maintain some variation, we introduced different vowel

contexts instead. The disyllables /pata/ and /poto/ are meaningful in Japanese (both are onomatopoeic expressions; "patapata" expresses the sound of footsteps, and "potopoto" the sound of dripping liquid), and therefore we decided to exclude /poto/ and its counterpart /topo/. However, we included /pata/ and /tapa/, because the vowel /a/ is the only one that allows a direct comparison with the previous study, and it allows the most open mouth configuration (MacNeilage et al., 2000).

### 3.3. Procedure

Participants were seated in front of a laptop computer (IBM ThinkPad X 40) connected to a USB microphone (Sony CARDIOID Dynamic Microphone F-V810). In each trial, participants were first presented a disyllable written in Japanese kana script, e.g. ?? /pata/ in black on white background in the middle of the screen. They were instructed to repeat the sequence presented, accelerating and decelerating in the rhythm of a visual timer they initiated by pressing the 'Space' key. The timer consisted of an alternation of black and white squares in the middle of the screen. It had a total duration of 16 s with an acceleration phase of 8 s, followed by a deceleration phase of the same length. The duration of presentation of each square started at 300 ms and gradually decreased until reaching 125 ms at 4 s, and 50 ms at 8 s. After that, durations again gradually increased symmetrically to acceleration. The timer was preceded by a blue square for 1000 ms. Participants were instructed to produce the first syllable on the black square, the second on the white, and so on. The visual timer did not have the function of precisely coordinating participants' production speed, but rather served as a global marker in order to decrease variability and to push participants to their limits. Participants were told that the timer would at one point reach an almost impossible speed, and that they should try to keep their production speed as fast as possible during that phase. Participants were encouraged to take a rest between trials whenever necessary, and there was a break between each block. There were six practice trials, during which the experimenter was present and made

sure participants had understood the instructions. Productions for each trial were recorded as separate sound files on the computer hard disk.

There were three experimental blocks during each of which the eight CVCV disyllables were presented once. Presentation order within each block was randomized independently.

### 3.4. Analysis

In order to assess if CVCV sequences would asymmetrically evolve into LC or CL CCV clusters, prosodic measurements based on vowel intensity were conducted. Analyses concentrated on the 3 s following the point of maximum acceleration, since Rochet-Capellan and Schwartz (2007) showed that articulatory asymmetries were most likely to occur in disyllable productions (hereafter, "utterances") of 300 ms or faster. 67% of utterances in the selected timespan fulfilled this premise ($M$ = 173 ms, min = 50 ms, max = 400 ms).

In general, the first production of each participant for one stimulus type was analyzed. If less than 50% of the participant's first production was codable for CV alternations (see below for exclusion criteria), the second production was chosen instead, and if this was still not codable, the third. A participant's second production was chosen in 4.6%, and the third in 0.6% of cases.

Intensity of each produced sequence was continuously estimated with the PRAAT software (Boersma & Weenink, 2009) using a 42.6 ms Kaiser-20 window with side-lobes below −190 decibel. Maxima and minima were automatically detected by consecutively searching time windows of 80 ms for their intensity maxima and minima from the beginning to the end of each 3-s sound file. The alternation of plosive consonants and vowels mostly led to clear minima and maxima in the resulting energy curves, with minima representing the complete closure in plosives, and maxima the vowel peaks. Manual parallel inspection of spectrograms and sound file ensured that no minimum and maximum value was missing or tagged twice.

Subsequently, the minima were manually labeled as either /p/ or /t/ by parallel inspection of spectrograms and sound files wherever possible. When a pattern evolved towards a CC cluster as /pt/ or /tp/ without any vowel peak in between, the corresponding minimum was labeled such. As the fast speed of some productions occasionally resulted in a deviation from the instructed voiceless /p/ or /t/, the labeling rule was that minima were labeled as /p/ or /t/ as long as a labial or coronal closure was clearly identifiable. This included voiced stops (/b,d/) or affricates (/p?/,/t?/), but excluded all other manners of articulation. Non-identifiable productions and speakers' errors, such as pauses, breathing, or repetition of the same CV sequence, were excluded from analysis. Overall, 12.6% of total productions were excluded this way.

As an index for articulatory asymmetry, the difference between the intensity scores of the vowel following a labial ($V_L$) and coronal ($V_C$) consonant was calculated for each utterance according to the formula: Delta $I$ = $I(V_C) − I(V_L)$. A Delta $I$ close to zero indicates a symmetrical utterance with similar intensities for the vowel after the labial and

coronal consonants, while a positive value indicates a tendency for /pt/ CC clusters, and a negative value for /tp/ CC clusters. Mean Delta $I$ values were obtained for each utterance, resulting in eight delta values for each of the 19 participants to be subjected to analysis. Among these, two utterances (/putu/ for one participant, and /tupu/ for another) only contained coronal consonants or non-identifiable productions and thus did not contribute any Delta $I$ value to the analysis.

### 3.5. Results and discussion

Fig. 1 plots Delta $I$ values for each stimulus type against utterance duration. Although most utterances center around zero, a visual inspection of the graphs shows that for utterances faster than 300 ms there are more positive than negative Delta $I$ values.

This asymmetry was statistically evaluated in two ways. First, a Chi-square goodness-of-fit test was conducted to compare frequencies of positive and negative Delta $I$ means,
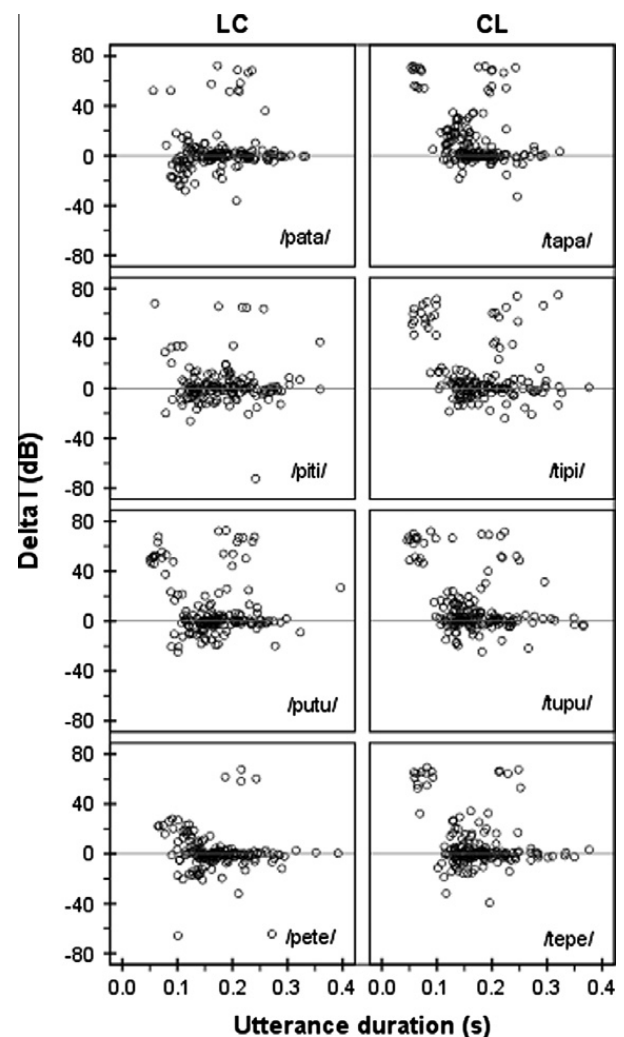


**Fig. 1.** $Y$ axis: Intensity variation between the vowel after the labial consonant and the vowel after the coronal consonant (Delta I). Positive values indicate evolvement towards a /pt/ cluster, negative values indicate evolvement towards a /tp/ cluster. $X$ axis: Duration of the respective utterance. Each circle represents one utterance.

showing that there were overall significantly more positive means than negative means (cf. Table 2). Second, a one-sample *t*-test was conducted to evaluate if Delta *I* means were significantly different from 0, showing that this was the case (cf. Table 2). To make sure that the lexical nature of /pata/ did not bias the results into the labial–coronal direction, the analyses were repeated after the exclusion of the disyllables /pata/ and /tapa/, which did not affect the direction of results (cf. Table 2). Lastly, analyses by token showed that Delta *I* is positive for all stimuli, with more positive than negative means for all stimuli (cf. Table 3; the value closer to zero is found for /pete/).

In summary, the current experiment shows higher articulatory stability of LC compared to CL plosive sequences in native Japanese adult speakers. This is despite the fact that in the lexicon of their native language, there are more CL plosive sequences than LC plosive sequences. As such, these results present strong evidence in support of an explanation of the LC bias based on articulatory factors. It is important to note, however, that an input-frequency explanation based on the overall LC bias in Japanese could have contributed to the LC production bias in the present experiment. Before further discussing the implications of these results, we first present an experiment exploring if and in which direction the perception of LC and CL plosive consonant sequences by Japanese speakers is biased.

## 4. Experiment 2: Perception in Japanese adults

To examine whether perception, like production, follows the universal tendency of an LC bias, or if the input frequency of the native language influences perceptual preferences, a verbal transformation task was conducted with Japanese participants. The experimental design and procedure closely resembled Sato et al. (2007). We focused on plosive consonants since the motor and perceptual explanations predict different outcomes (LC versus CL biases respectively), and decided to present each participant with stimuli recorded by a Japanese speaker and by a French speaker in order to determine potential effects of the phonological/phonetic properties of the stimuli.

### 4.1. Participants

Sixteen students and university staff (seven females) of several universities in Tokyo (mean age: 24.2 years; range: 20–38) with no speaking or hearing problems participated in the experiment for payment. They were all native speakers of Japanese. Due to a program error, four additional participants were tested but they were not presented with the full set of stimuli. The data from these participants were therefore excluded from analysis.

### 4.2. Stimuli

The target sequences used here were the same as in Experiment 1, excluding the vowel context /a/: three LC sequences (/pete/,/piti/,/putu/) and their CL counterparts (/tepe/,/tipi/,/tupu/). These sequences differ from those in Sato et al. (2007) in two ways. First, instead of presenting sequences in the vowel contexts /a,i,o/, we chose the vowel contexts /e,i,u/, because /pata/ and /poto/ are lexical in Japanese (cf. Section 3.2). Note that while in the production study, we did not exclude /a/ in order to keep one stimulus constant with the previous study in French, we could exclude it in the present perception study since the vowel context /i/ was used by Sato et al. (2006). As a second change, we presented stimuli recorded by native speakers of two languages: a male native speaker of Tokyo Japanese, and a male native speaker of metropolitan French.

In order to obtain the stimuli, several tokens of the CV sequences /pe/, /pi/, /pu/, /te/, /ti/, /tu/ were recorded in isolation in a soundproof room. Both speakers were instructed to pronounce CV sequences at a natural conversation rate while keeping an even intonation and intensity. The items were digitized on the hard disk of a computer at a 44.1 kHz sampling rate. Then, for each language and for each vowel context, one *p*-initial and one *t*-initial CV sequence (e.g., /pe/ and /te/ in French) were selected to form one token pair. Consonant and vowel duration, mean consonant and vowel intensity, F1, F2, and F3 formant values, as well as minimum, maximum, and mean vowel pitch were matched as closely as possible within each token pair (Table 4). From each of these token pairs, two experimental stimulus file were constructed, both consisting of 300 alternated repetitions of the two syllables, one starting with the *p*- and the other one with the *t*-inital CV sequence. This resulted in a total of twelve stimulus files (3 vowel contexts × 2 initial CVs × 2 voices).

In order to reflect the silent period before stop release, a 75 ms pause preceded each CV sequence. On average, Japanese token pairs were 537 ms long, and French token pairs were on average 546 ms long.

### 4.3. Apparatus and procedure

Participants were individually seated in front of a laptop computer (IBM ThinkPad X 40) in a sound-attenuated room and presented the experimental stimuli binaurally via a pair of headphones (audio-technica ATH-A 500) at a comfortable

**Table 2**
Mean Delta *I* values, number of positive and negative means, and statistical analyses for disyllables overall and under exclusion of the /a/ vowel context. $\chi^2$–test compared the number of negative and positive means. A positive Delta *I* value indicates evolvement towards an LC cluster. One-tailed *t*-test compared mean Delta *I* against 0.

| | Mean Delta *I* | Negative means | Positive means | $\chi^2$ – Test | | *t*-Test | |
|---|---|---|---|---|---|---|---|
| | | | | $\chi^2$ (df) | *p* | *t* (df) | *p* |
| All disyllables | 5.20 | 54 | 96 | 11.76 | .001 | 4.78 (149) | <.001 |
| Disyllables excluding pata/tapa | 4.68 | 40 | 72 | 9.14 | .002 | 3.71 (111) | .001 |

**Table 3**
Mean Delta *I* values and number of positive and negative means by token. Positive Delta *I* values indicate evolvement towards an LC cluster, and negative means indicate evolvement towards a CL cluster.

|  | pata | tapa | piti | tipi | putu | tupu | pete | tepe |
|---|---|---|---|---|---|---|---|---|
| Mean Delta *I* | 4.05 | 9.42 | 1.07 | 8.49 | 5.80 | 8.65 | 0.31 | 4.04 |
| Negative means | 8 | 6 | 9 | 6 | 6 | 4 | 9 | 6 |
| Positive means | 11 | 13 | 10 | 13 | 12 | 14 | 10 | 13 |

**Table 4**
Acoustic properties of presented CV syllables. Acoustic properties were matched as closely as possible for each CVCV pair. Consonant duration reflects voice onset time (VOT) plus the added 75 ms of silence.

|  |  | /pe/ | /te/ | /pi/ | /ti/ | /pu/ | /tu/ |
|---|---|---|---|---|---|---|---|
|  |  | *Japanese* |  |  |  |  |  |
| Duration (ms) | Consonant | 101 | 112 | 131 | 134 | 105 | 121 |
|  | vowel | 156 | 159 | 158 | 156 | 137 | 141 |
| Intensity (dB) | VOT | 71 | 71 | 63 | 68 | 71 | 72 |
|  | vowel | 77 | 77 | 75 | 77 | 77 | 80 |
| Vowel formant (Hz) | $F_1$ | 577 | 577 | 376 | 375 | 335 | 376 |
|  | $F_2$ | 2189 | 2269 | 2471 | 2471 | 1786 | 1786 |
|  | $F_3$ | 2793 | 2793 | 3398 | 3317 | 2471 | 2552 |
| Vowel pitch (Hz) | Min | 90 | 90 | 100 | 103 | 91 | 93 |
|  | Max | 139 | 146 | 136 | 150 | 126 | 127 |
|  | Mean | 110 | 110 | 111 | 112 | 104 | 108 |
|  |  | *French* |  |  |  |  |  |
| Duration (ms) | Consonant | 104 | 120 | 124 | 145 | 119 | 128 |
|  | vowel | 143 | 158 | 133 | 160 | 148 | 156 |
| Intensity (dB) | VOT | 68 | 64 | 63 | 66 | 67 | 65 |
|  | vowel | 72 | 72 | 73 | 73 | 72 | 73 |
| Vowel formant (Hz) | $F_1$ | 385 | 344 | 324 | 283 | 335 | 375 |
|  | $F_2$ | 1945 | 1924 | 2026 | 2046 | 1061 | 1141 |
|  | $F_3$ | 2613 | 2573 | 2876 | 2795 | 2753 | 2713 |
| Vowel pitch (Hz) | Min | 100 | 93 | 98 | 110 | 100 | 99 |
|  | Max | 117 | 117 | 130 | 116 | 121 | 136 |
|  | Mean | 107 | 103 | 113 | 113 | 111 | 108 |

sound level. Different from Sato et al. (2007), participants did not respond orally as soon as they perceived a change, but pressed response keys instead. This procedure was chosen, because in contrast to previous studies on verbal transformations (Pitt & Shoaf, 2002; Sato et al., 2007; Warren, 1961), which had the additional purpose of exploring the space of possible transformations, the current study was solely interested in the ratio of LC–CL perception. As such, oral responses including the exact nature of each transformation were not necessary.

On a QWERTY laptop keyboard, the "I" and "O" keys were covered with stickers clearly labeled as "P" and "T". Participants were asked to press the left key with the index finger, and the right with the middle finger of their right hand. The labeling of keys was counterbalanced across participants, so that for half of the participants "P" was left of "T", and "T" was left from "P" for the other half. The "G" key was covered with an unlabeled blue sticker.

Participants were first introduced to the phenomenon of verbal transformations by listening to a repeated sequence of either the disyllable /mono/ or /nomo/ (counterbalanced across participants). After listening, they were asked if they had perceived any change in the sequence, and in case not, they were explained that their perception of the sequence might change from /mono/ to /nomo/, or vice versa, during listening. In the subsequent practice trial, they listened to the same sequence for about 1 min, and were instructed to press response keys as follows. They were asked to initiate the trial by pressing the space key, and to press either the "N" or "M" key as soon as the sound sequence had started in order to indicate whether they had perceived /n/ or /m/ at the beginning of the sequence. Subsequently, they were asked to press the response keys only if perceiving a change. If they perceived a change from /mono/ to /nomo/, they were asked to press "N", and vice versa. It was emphasized that they might not perceive any change, or else very frequent changes from time to time. They were also told that they might perceive a change to a completely different sequence including neither /m/ nor /n/, and to press the blue key in this case. This option was included based on the findings of Sato, Vallee, Schwartz, and Rousset (2007; Experiment 1) that a change to a different sequence was perceived in 31% of trials.

At the beginning of each trial, a fixation cross appeared for 500 ms, followed by a blank screen for the duration of the sound sequence. After the practice trial, participants were explained that sequences would now start with either /p/ or /t/ instead of /m/ or /n/. After each of the twelve trials,

a screen informed participants about the number of trials completed and encouraged them to rest as long as necessary. Trial order was randomized across participants. The experiment was run with E-Prime 2.0, and answers were saved on the hard disc of the computer.

### 4.4. Results and discussion

Percentages of LC, CL and other responses for the 12 different experiment files are presented in Fig. 2. Perceptual stability of each stimulus was obtained by summing up the time the initial disyllable was perceived (LC for p-initial sequences and CL for t-initial sequences) as indicated by button presses. Then, relative perceptual stability (our dependent variable) was obtained by dividing the perceptual stability of each stimulus by the total time the stimulus was perceived as either LC or CL. Overall, the time spent perceiving a stimulus as either an LC or a CL pattern was 90.5% for Japanese sequences, and 90.4% for French sequences, showing that for stimuli of both languages, the recognition rate was similarly high.

A three-way repeated-measures ANOVA with the within-subject factors of Sequence (LC versus CL), Language Presented (Japanese versus French) and Vowel (e, i, u) was performed. Bonferroni-corrected post hoc paired comparisons at a significance level of $p < .05$ were conducted where appropriate. This analysis revealed a significant main effect of Sequence [$F(1, 15) = 6.02$, $p = .027$, $\eta_p^2 = .286$], indicating higher perceptual stability for CL sequences ($M = .57, SD = .11$) than for LC sequences ($M = .50, SD = .07$). There was also a significant interaction between Sequence and Language Presented [$F(1, 15) = 8.232$, $p = .012$, $\eta_p^2 = .354$]. Post-hoc paired comparisons showed that the effect of Sequence was significant for Japanese stimuli ($p = .004$; stability of CL sequences: $M = .62$, SD = .13; stability of LC sequences: $M = .44$, SD = .12), but not for French stimuli ($p = .336$; stability of CL sequences: $M = .52$, SD = .12; stability of LC sequences: $M = .56$, SD = .12). This indicates a CL bias for Japanese stimuli, and no bias for French stimuli.

The results show an overall CL bias for native speakers of Japanese, which is congruent with perception-based predictions since there is a CL bias for plosive sequences in the Japanese lexicon. These results complement the previous results in French (Sato et al., 2007), in which an LC bias was found for plosive sequences that exhibit an LC bias in the French lexicon. This suggests effects based on input properties. However, the effect observed in our experiment was influenced by language of presentation, such that a CL bias occurred for Japanese stimuli, but not for French stimuli. Since it was unclear whether this language effect was due to some idiosyncratic properties of the stimuli recorded, or whether they reflected language-specific processing effects, we decided to replicate Experiment 2 with French participants.

## 5. Experiment 3: Perception in French adults

Experiment 3 tested perceptual biases in French participants with the same stimuli as those presented to Japanese participants in Experiment 2.

### 5.1. Participants

Sixteen students and university staff (12 females) of Université Paris Descartes (mean age: 26.3 years; range: 22–44) with no speaking or hearing problems participated in the experiment for payment. They were all native speakers of French.

### 5.2. Stimuli

The stimuli were the same as in Experiment 2. As the stimuli were primarily constructed for the study of Japanese participants, it could not be avoided that of these, /pete/, /piti/, and /tipi/ were very low frequency words in French. The word corresponding to /pete/ is "péter", meaning "to fart", with a frequency of 17.09, the word corresponding to /piti/ is "Pythie," the Greek oracle, with a frequency of .54, and the word corresponding to /tipi/ the native American tent "teepee", with a frequency of .01. Frequencies are according to counts in Lexique.org (New, Pallier, Ferrand, & Matos, 2001).

### 5.3. Procedure

The procedure was the same as in Experiment 2.

### 5.4. Results and discussion

Percentages of LC, CL and other responses for the 12 different experiment files are presented in Fig. 3. The time spent perceiving a stimulus as either LC or CL was 80.0% for Japanese sequences, and 83.1% for French sequences.

As in Experiment 2, a 3-way repeated-measures ANOVA with the within-subject factors of Sequence, Language Presented and Vowel was performed. Due to a violation of the sphericity assumption in the interaction between Sequence and Vowel, Greenhouse–Geissner corrected values are reported for this interaction. There was no main effect of Sequence [$F(1, 15) = 0.41$, $p = .531$, $\eta_p^2 = .027$; stability of CL sequences: $M = .56$, SD = .12; stability of LC sequences: $M = .58$, SD = .10]. However, there was a significant interaction between Sequence and Language Presented [$F(1, 15) = 7.93$, $p = .013$, $\eta_p^2 = .346$]. Post-hoc paired comparisons showed a significant LC bias for French stimuli ($p = .008$; stability of CL sequences: $M = .49$, SD = .15; stability of LC sequences: $M = .63$, SD = .13), but no significant effects for Japanese stimuli ($p = .119$; stability of CL sequences: $M = .63$, SD = .15; stability of LC sequences: $M = .53$, SD = .17). There was also a significant interaction between Sequence and Vowel [$F(1.47, 21.97) = 5.23$, $p = .021$, $\eta_p^2 = .258$]. The difference in stability between sequences was significantly different for the vowels /u/ ($p = .037$; stability of CL sequences: $M = .64$, SD = .14; stability of LC sequences: $M = .52$, SD = .16) and /i/ ($p = .020$; stability of CL sequences: $M = .49$, SD = .22; stability of LC sequences: $M = .63$, SD = .14), with a CL bias for the former, and an LC bias for the latter.

In summary, although French participants did not show a main effect of sequence, this lack of an overall effect is due to the fact that we presented stimuli recorded in either French or Japanese, the former giving rise to an LC bias
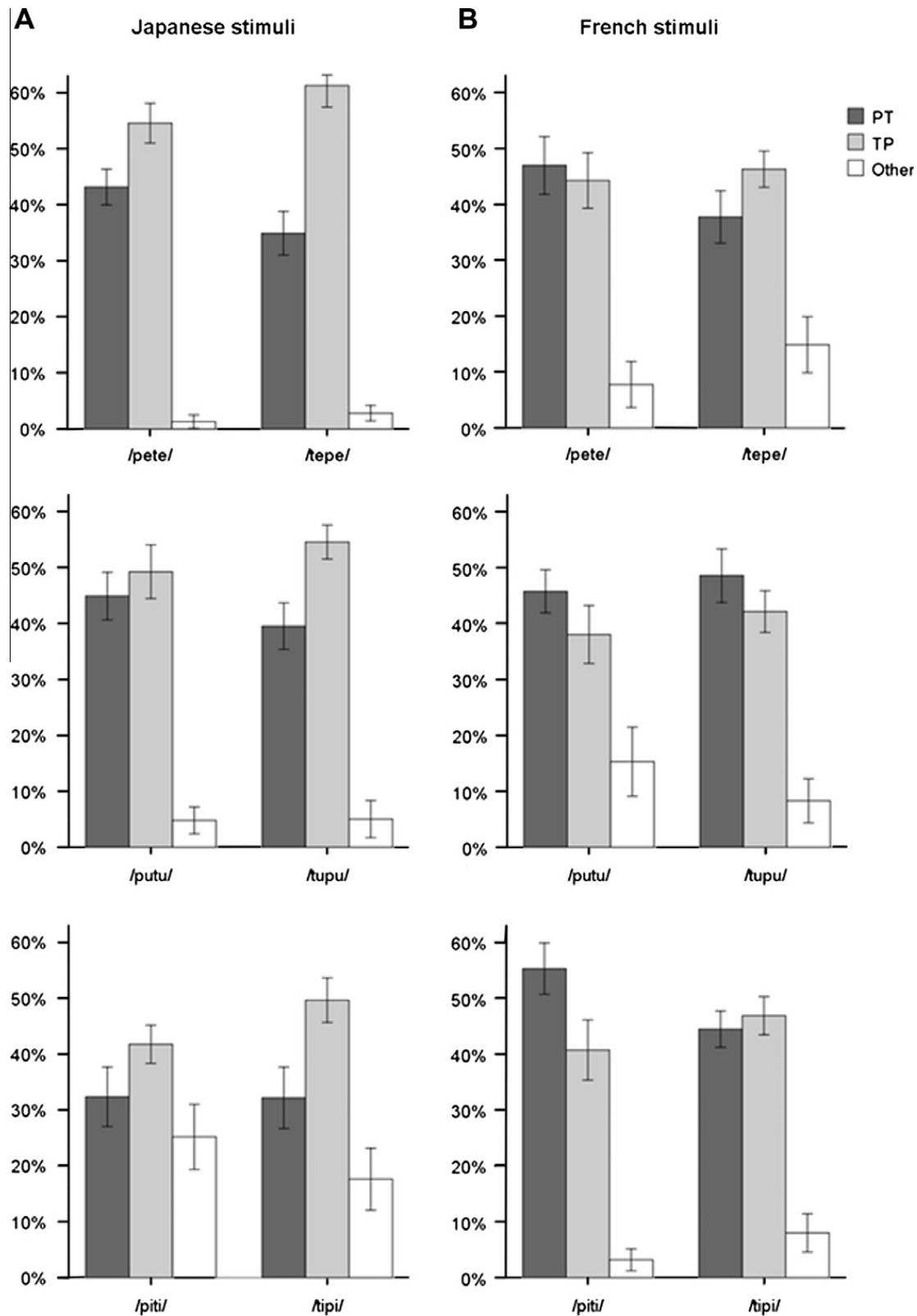
**Fig. 2.** Mean ratios of perception as LC or CL sequence for Japanese participants. (A) Stimuli presented in Japanese and (B) Stimuli presented in French. Error bars represent ±1 SE.

and the latter giving rise to no bias. The LC effect for native language stimuli thus replicates the previous findings (Sato et al., 2007), extending them to new sequences and new recordings.

Interestingly, both the vowel context and language of presentation mattered for French participants. The former was not found in Experiment 2 for Japanese participants. It is unclear why there was such an unpredicted CL bias
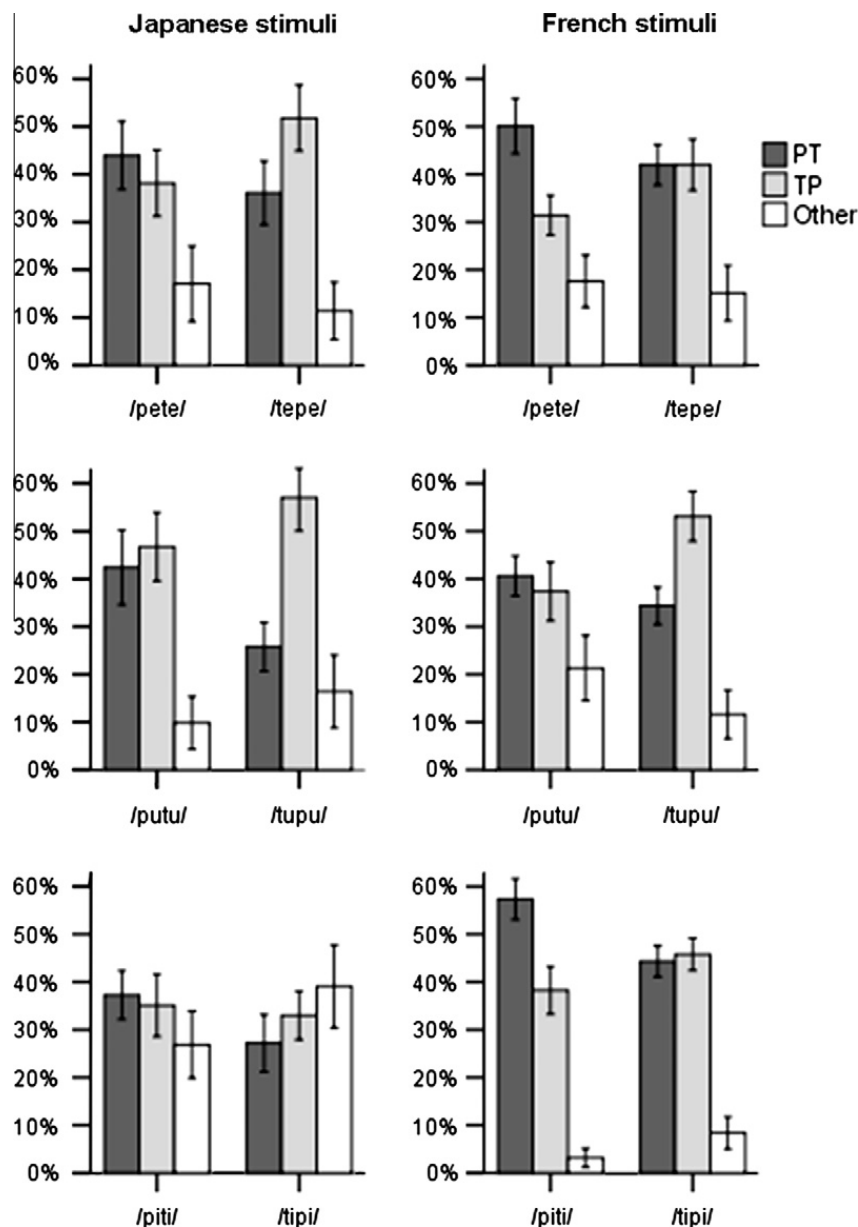
**Fig. 3.** Mean ratios of perception as LC or CL sequence for French participants. (A) Stimuli presented in Japanese and (B) Stimuli presented in French. Error bars represent ±1 SE.

for the /u/ vowel context, though a possible explanation comes from informal observations given after the task by some participants, who declared having perceived "tu peux" (you can), probably as a misperception of the /tupu/ and /putu/ sequences. This misperception, which appears slightly larger for the Japanese stimuli, might have been favored by phonetic properties of the /u/ vowel in Japanese, which is pronounced with compressed lips, is unrounded but without spreading (Okada, 1991). As such, it is different in realization from the French /u/ (Vance, 1987).

As for the language of presentation effect, French participants' perceptual bias turned out exactly opposite from that of Japanese participants, with an LC bias for native stimuli only. Therefore, our findings reveal a crossed result with a perceptual bias consistent with native language input for both Japanese and French listeners, thus a CL bias for Japanese participants and an LC bias for French partic-

pants, but with native stimuli only. This clearly establishes linguistic input influences on perceptual biases.

However, our findings also raise the question of why these respective biases are not found with non-native stimuli. A first possibility is that there is a modulation of the listeners' native language biases by some idiosyncratic characteristics of either the French or Japanese stimuli used.[5] Although the stimuli were matched as well as possible on their acoustic properties, there was one consistent difference between labial-initial and coronal-initial sequences in the French, but not in the Japanese, stimuli: The length of the vowels following /p/ was always shorter than the length of the vowels following /t/ (see Table 5). We had not been

---

[5] We thank Jean-Luc Schwartz for raising this issue during the review process.

able to neutralize this difference in our original stimuli because the stimuli were selected carefully from the recordings of a native speaker of French who produced these differences systematically. It might have been a property of our speaker, or a property of French. A posteriori, combined with the language effect (CL bias in Japanese speakers, LC bias in French speakers), this length difference in the French stimuli could potentially explain part of our results. If the length difference in the French stimuli induced more LC responses in Japanese participants, this could explain why the CL bias that was found for Japanese speakers listening to Japanese stimuli was neutralized when they listened to the French stimuli. It is also possible that the French participants' LC bias was enhanced by the durational differences of the French stimuli, while it was comparatively reduced when listening to the Japanese stimuli in this within-subject design.

To test the possibility that the durational differences in the French stimuli contributed somehow to the results of Experiments 2 and 3, a first control experiment, Experiment 4a, tested Japanese participants' perceptual biases with modified French stimuli that were matched on vowel length. If participants show a CL effect with this altered material, this would suggest that the differential vowel length could indeed have been a reason for the previous absence of a bias in response to French stimuli. On the other hand, if they show effects comparable to those of Experiment 2, this would suggest that the reason for the pattern found in Experiments 2 and 3 has to be found elsewhere.

An alternative explanation for the pattern of results found is that listeners process native and non-native language stimuli differently, applying their perceptual biases only to the former. In order for listeners to do so, they must – implicitly or explicitly – be able to tease apart the native and non-native stimuli. Their ability to do so was explored in one further control experiment, Experiment 4b, which assessed whether Japanese participants explicitly can tell the native stimuli apart from the non-native ones.

## 6. Experiment 4

Experiment 4 was designed to take a closer look at the language of presentation effect found in Experiments 2 and 3. Experiment 4a assessed how the length difference properties of the original French stimuli might have modulated the biases observed. Experiment 4b evaluated the extent to which participants can tell apart native and non-native stimuli. These control experiments were run with Japanese participants, since they are the main focus of the present study. The same participants ran both control experiments, always in the order Experiment 4a, then 4b.

### 6.1. Experiment 4a

In the original French stimuli, the length of the vowels following /t/ was always longer than the length of the vowels following /p/. We have discussed above that this is one possible reason for the language of presentation effects found in Experiments 2 and 3, since this difference might have induced more LC responses, resulting in a null effect for the Japanese adults listening to the French stimuli. Therefore, we reasoned that if this length difference is neutralized, then a CL bias should emerge in Japanese participants.

#### 6.1.1. Participants

Sixteen students (six females) of several universities in Tokyo (mean age: 20.6 years; range: 18–29) with no speaking or hearing problems participated in the experiment for payment. They were all native speakers of Japanese.

#### 6.1.2. Stimuli

The stimuli consisted of modified versions of the French stimuli from Experiments 2 and 3. In the original stimuli, the length of the vowels following /t/ was consistently longer than the length of the vowels following /p/. Note that /t/s were also systematically longer than /p/s, but this was also the case for the Japanese stimuli and therefore not considered a relevant factor differentiating between the French and Japanese stimuli (cf. Table 4). In order to match vowel durations, the vowel length of the vowels following /t/ was shortened to match the length of the vowels following /p/ pairwise for each vowel context (/i/, /e/, and /u/). Vowels were shortened by removing a part from the steady state middle section of each vowel. Resulting vowel lengths are given in Table 5.

#### 6.1.3. Procedure

The procedure was the same as in Experiment 2, except that instead of 12 trials (six French, six Japanese), participants were only presented with six French trials.

#### 6.1.4. Results and discussion

The overall time of perceiving the sequence as either LC or CL was 75%. A repeated-measures ANOVA with the within-subject factors Sequence and Vowel revealed no significant main effect for sequence [$F(1,15) = 0.18$, $p = .676$, $\eta_p^2 = .012$; stability of CL sequences: $M = .59$, SD = .16; stability of LC sequences: $M = .62$, SD = .20], a marginally significant effect of vowel [$F(2,30) = 3.00$, $p = .065$, $\eta_p^2 = .167$; /e/: $M = .58$, SD = .19; /u/: $M = .69$, SD = .20; /i/: $M = .56$, SD = .12], and no interaction [$F(2,30) = 1.12$, $p = .34$, $\eta_p^2 = .069$]. A follow-up on the marginal vowel effect revealed no significant differences between any of the three possible pairings of vowels (for the comparison /e/–/u/: $t = 1.64$, $p = .105$; for /i/–/e/: $t = 0.57$, $p = .573$; for /u/–/i/ $t = -1.250$, $p = .215$).

The manipulation of vowel length thus did not affect the perceptual bias in Japanese participants, who still show no bias when presented with the modified French stimuli. Therefore, we can conclude that the vowel length difference in the original French stimuli, which might have favored LC responses, were not the reason for the absence of a CL effect

**Table 5**
Length of original (used in Experiments 2 and 3) and shortened (used in Experiment 4a) vowels, French stimuli.

| | | /pe/ | /te/ | /pi/ | /ti/ | /pu/ | /tu/ |
|---|---|---|---|---|---|---|---|
| Duration (ms) | Consonant | 104 | 120 | 124 | 145 | 119 | 128 |
| | Vowel old | 143 | *158* | 133 | *160* | 148 | *156* |
| | Vowel new | | 140 | | 133 | | 147 |

for non-native stimuli in Japanese participants. There was no comparable difference in /p/-initial and /t/-initial sequences in the Japanese material that could be manipulated and tested on French participants. At this point, we turn to the second possible explanation for the language of presentation effect that we discussed, namely that native and foreign language stimuli might be processed in different ways. We explore a precondition for a difference in the processing of native and non-native stimuli, which is that adults are able to explicitly or implicitly recognize which of our stimuli are in their native language and which are not.

### 6.2. Experiment 4b

In order to test explicit identification of native and non-native stimulus material, Japanese participants were presented the CV or CVCV sequences that constituted the original stimuli used in Experiments 2 and 3, and were asked to decide whether they heard a Japanese or a foreign speech sound.

#### 6.2.1. Participants

Participants were the same as in Experiment 4a.

#### 6.2.2. Stimuli

Stimuli consisted of the CV or CVCV sequences constituting the original stimuli. In the first block, CV sequences, for instance /pe/, were presented, and in the second block, CVCV sequences, for instance /pete/, were presented. Each of the 12 CV and 12 CVCV sequences that were used in Experiments 2 and 3 were presented once to each participant.

#### 6.2.3. Procedure

Participants were seated individually in front of a laptop computer in a sound-attenuated room and presented the experimental stimuli binaurally via a pair of headphones. Preceding the first block, they were explained that they were going to hear speech sounds of the length of one kana symbol, and were instructed to respond by button press with the index finger of the right hand if it was a 'Japanese' sound, and the index finger of the left hand if it was a 'foreign language' speech sound. Preceding the second block, instructions informed them that they would now hear sequences of two kana symbols.

#### 6.2.4. Results and discussion

Participants' mean 'Japanese' responses were taken as the dependent variable.

A two-way repeated measures ANOVA with the factors Language Presented (Japanese, French) and Sound Type (CV, CVCV) revealed a main effect of Language Presented [$F(1, 15) = 22.73$, $p < .001$, $\eta_p^2 = .60$; 'Japanese' responses for Japanese stimuli: $M = .55$, SD = .15; 'Japanese' responses for French stimuli: $M = .25$, SD = .17], but no effect of Sound Type [$F(1, 15) = .30$, $p = .59$, $\eta_p^2 = .02$], and no interaction [$F(1, 15) = 2.37$, $p = .14$, $\eta_p^2 = .14$] (cf. Fig. 4).

Japanese participants were able to tell apart the foreign from the native language stimuli, even when they only heard a CV sequence. This ability is a precondition for participants to process native and foreign language stimuli dif-
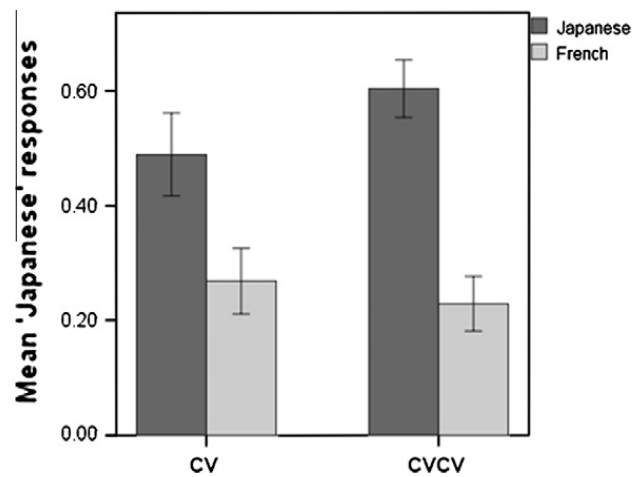


**Fig. 4.** Mean 'Japanese' responses for the Japanese (dark grey) and French (light) grey CV (left) or CVCV (right) sequences that constituted the experimental stimuli of Experiments 2 and 3.

ferently and, hence, to possibly produce the language of presentation effects observed in Experiments 2 and 3. In the General Discussion, we will come back to this issue and discuss how these perceived differences might affect the expression of the LC and CL perceptual biases.

### 7. General discussion

In the present study, we investigated the question of whether the LC bias is determined by motor factors, perceptual factors, or both. Previous studies have found evidence for both motor (Rochet-Capellan & Schwartz, 2007) and perceptual (Sato et al., 2007) influences, the former indicating a general mechanism, and the latter a language-specific one. However, due to the fact that these studies were exclusively conducted in languages with a higher frequency of LC compared to CL sequences, they have not been able to isolate the relative influence of perceptual input on both productive and perceptual preferences.

Japanese has been claimed to be a language with the opposite bias (MacNeilage et al., 1999), making it a candidate language for disentangling these accounts. Due to the fact that this claim was based on a very small sample of words, we conducted a large-scale corpus analysis in order to reevaluate these previous findings. Across corpora and analyses, we found that the subset of plosives consistently showed a CL bias, while the subset of nasals, as well as the analysis of all segments, showed an LC bias. The deviation of our current findings from the previous ones is possibly due to the small sample size, as well as the very selective vocabulary covered in the travel dictionary used in the previous study. Finding an overall LC bias in the only language that has been claimed to favor the opposite pattern to date leads further support to the notion that the LC bias is predominant in languages of the world (MacNeilage & Davis, 2000). Importantly, however, despite the overall LC bias, the subset of plosives did show a CL bias. Plosives are among the first segments children produce (cf. MacNeilage et al., 2000), which makes an investigation of this subset of special interest to early speech develop-

ment: If there is a universal tendency to produce plosives from early on, does the sequencing of plosive consonants also follow such a tendency, i.e., would Japanese infants also show an early LC bias in production? Our adult production data would suggest they do, but the infant literature still lacks data from languages with the opposite bias. Likewise, data from French infants (Gonzalez-Gomez & Nazzi, 2012b) suggest that the perceptual LC bias is an acquired one, and it will be of interest to determine if Japanese infants develop a CL bias for plosives at a similar age.

Having singled out a subset of segments with a consistent CL bias in Japanese, we investigated the productive preferences of Japanese adults with regard to plosive LC and CL sequences. In the context of a speeded articulation task, Japanese participants, like French adults (Rochet-Capellan & Schwartz, 2007), showed a tendency to reduce LC and CL CVCV clusters into LC CCV clusters. This suggests that LC plosive sequences are articulatory more stable for speakers of Japanese despite the higher frequency of CL plosive sequences in their input, and provides strong support for an account that bases the higher prevalence of LC patterns in languages of the world on characteristics of the human motor system. Furthermore, this result seems remarkable in the light of Japanese phonotactics, in which CCV consonant clusters are illegal. However, Tokyo Japanese entails phonological devoicing after the vowels /u/ and /i/, which in fact regularly leads to the production of consonant clusters (e.g., "tsukuru" → "ts'kuru"). Moreover, work on the perception of CCV clusters (Dupoux, Kakehi, Hirose, Pallier, & Mehler, 1999) has shown that Japanese listeners, when presented with consonant clusters, perceive epenthetic vowels (e.g., /ebzo/ is perceived as /ebuzo/), illustrating that they have a repair mechanism for devoiced forms to fit into native phonology. Thus, our data might be a nice illustration of the fact that the production system is capable of producing CCV clusters, and is even inclined to do so if it benefits articulatory ease, while the perception system provides a mechanism to fit ill-formed sequences into native language phonology.

Contrary to the results in production, Japanese listeners showed a language-specific bias in online speech perception and preferred CL over LC plosive sequences. These findings are in line with numerous studies showing an influence of native language phonology on segmentation (e.g., McQueen, 1998; Mersad & Nazzi, 2011; Pena et al., 2002; Saffran et al., 1996; Weber & Cutler, 2006). They are also consistent with the developmental finding that French infants start out without any bias at 6 months of age, but develop an LC bias by 10 months (Gonzalez-Gomez & Nazzi, 2012b; Nazzi et al., 2009), which is in line with other studies of infant speech perception (e.g., Friederici & Wessels, 1993; Jusczyk et al., 1993; Jusczyk et al., 1994). The finding that Japanese listeners show a perceptual CL bias despite the fact that their overall input (i.e., considering all manners of articulation) is biased towards LC also suggests that this perceptual bias applies at the level of manner of articulation rather than overall. These results complement recent findings with French adults (Sato et al., 2007) and infants (Gonzalez-Gomez & Nazzi, 2012a,b). French has an LC bias both overall and for the subgroup of plosives; however, the subgroup of fricatives shows a CL bias. Gonzalez-Gomez

and Nazzi (2012a) found that 10-month-olds' perceptual bias for different manners of articulation is directly related to the input bias in the respective manner, i.e. LC for plosives and nasals, but CL for fricatives. Furthermore, voiced plosives have a CL bias, but nevertheless French adults show an LC bias for this subgroup (Sato et al., 2007), which is however compatible with the overall LC bias at the level of all plosives. These findings in combination with the results of the present study suggest a picture of the perceptual LC and CL biases applying at the level of manner of articulation. Further studies will be needed to clarify if Japanese listeners indeed show a perceptual LC bias if presented with stimuli in other manners of articulation. Due to the lack of sufficient labial segments in fricatives, this study would have to be conducted on nasals.

Both Japanese and French participants show an influence of the language presented. Japanese listeners exhibited a CL bias for the Japanese stimuli only, while French listeners in turn showed an LC bias exclusively for the French stimuli. Thus, both groups of listeners do not show a statistically significant bias when listening to their non-native language stimuli, indicating that this language-specific sequential bias is likely not generalized to instances in other languages. In two control experiments, we explored two explanations for this pattern of results.

According to the first explanation (our least preferred), the pattern could have been an artifact of an acoustic property of our stimuli, namely the fact that there was a difference in vowel length for vowels after /t/ and /p/ in the French stimuli, which was not present in the Japanese stimuli. If this crosslinguistic difference induces more LC responses by the French compared to the Japanese stimuli, then coupled with the different biases present in Japanese (CL bias) and French (LC bias) adults, we would observe the pattern found in Experiments 2 and 3. However, when this difference was controlled for and the new stimuli presented to Japanese adults (Experiment 4a), the results failed to reveal a bias with these controlled non-native French stimuli. This suggests that this property of the original French stimuli is unlikely to explain the pattern of findings.

According to a second explanation, the pattern could also be explained by the presence of different biases in Japanese (CL bias) and French (LC bias) adults, which would only be applied to native language stimuli due to the fact that adults would process native and foreign language stimuli differently. Although proving this explanation was beyond the scope of the present study, we at least tested a prerequisite for this explanation in Experiment 4b, namely that Japanese adults can distinguish between the Japanese and French stimuli used in Experiments 2 and 3. The results established that this was the case, implying that Japanese adults could process these two groups of stimuli differently. In the following, we discuss the factors that might determine these differences in processing our native and non-native stimuli, and implications on the level at which the LC and CL biases might operate.

If the LC and CL biases are phonological in nature, as we suggested, one would expect them to generalize to the same phoneme string across different acoustic/phonetic realizations, and possibly even across languages. This was, however, not the case in the current study, suggesting the

influence of the acoustic/phonetic properties of the stimuli. If the acoustic/phonetic explanation is to be sufficient, it would predict similar perception of the same stimuli across listeners. Yet, Experiments 2 and 3 demonstrated clear differences between Japanese and French listeners' performance. Moreover, Experiments 2 and 4a taken together showed similar performance of Japanese adults for the original and the acoustically modified French stimuli.

Another way in which acoustic/phonetic differences might influence the expression of the biases is by interfering with how phonological information is retrieved from the stimuli. One extreme version of such interference would be that it would block the encoding of the information at the phonological level. This does not seem to be what is happening in the present study, since this would have probably led to very different perceptions of the native and foreign stimuli, and important difficulties in processing the non-native stimuli. However, no participant reported hearing some stimuli that they had trouble processing, nor did we observe differences in the amount of time stimuli in both languages were reported to be perceived as LC or CL. A less extreme version is that non-native realizations add noise to the acoustic-to-phonological encoding processes. Indeed, although the Japanese and French stimuli correspond to the same phonemes, these phonemes are realized differently in the two languages at the acoustic/phonetic level. For example, as mentioned earlier, the /u/ vowel is pronounced with compressed lips, is unrounded but without spreading in Japanese (Okada, 1991), which is not the case for the rounded French /u/ (Vance, 1987). Moreover, plosives in Japanese can be weakly aspirated (Okada, 1991), with voice onset time (VOT) of Japanese voiceless stops falling between average VOTs for unaspirated and aspirated stops in other languages (Riney, Takagi, Ota, & Uchida, 2007), while plosives in French are mostly unaspirated (Fougeron & Smith, 1993). Future studies will be needed to establish the impact of acoustic/phonetic variability on the expression of the LC and CL biases, both for cross-linguistic variability but also for within-language variability. No matter the outcome, it remains that our study clearly reveals differences in how sequences of alternating plosive labial- and coronal-initial syllables are organized at the perceptual level (LC bias in French, CL bias in Japanese).

At this point, we would like to discuss the fact that the findings of the present study are difficult to reconcile with a perception–action link in determining the LC bias. In their study of the relation between speeded production and verbal transformations, Sato et al. (2007) found support for the notion of a perception–action link in the LC bias, suggesting that this link plays a role in the case of the perceptual LC bias in French. The results of the current study do not exclude this possibility for French; however, the observed dissociation between perception and action for Japanese suggests that, if such a perception–action link is present, other factors can override it. In other words, the Japanese data suggest that, when there is a CL bias in the input, it wins over the production constraint for an LC bias.

Remaining questions are, firstly, how prevalent the LC bias actually is in languages of the world, and where the plosive CL bias in Japanese originates. Although the corpora examined by MacNeilage et al. (1999) cover several lan-

guage families, they are far from complete. Historically, Japanese has borrowed heavily from the Chinese language in both script and sound, and although controversial, some roots in the Korean language are also assumed (Lee & Hasegawa, 2011). Starting out with languages that are close to Japanese, further languages have to be examined in order to get a better picture of the pervasiveness of the LC bias across languages of the world.

Secondly, adult listeners' biases for different subclasses of consonants in different languages are of interest. Both our results and Gonzalez-Gomez and Nazzi (2012a) suggest that listeners develop input-specific biases at the level of manner of articulation. Exploring further subclasses of consonants in different languages with different predictions for different manners of articulation will be necessary in order to confirm this tendency.

Thirdly, in light of the adult findings, it is of interest to evaluate what we can expect with regard to infants' developing production and perception. With regard to perception, both the findings of the current study and the findings with French infants allow the prediction that Japanese infants will show an input bias, i.e. a CL bias for plosives and an LC bias otherwise. With regard to infants' early productions, if we assume that articulatory stability plays a major role, we can expect an LC bias as found in Japanese adults. However, other factors might influence early productions, especially early words. MacNeilage and Davis (2000) found an LC bias in infants' first words, but not yet in babbling. As infants have been exposed to their native language's input for quite some time by the time they start producing words, an influence of input frequency cannot be excluded based on these data. Direct support for such an influence on the development of an LC bias in production comes from Fikkert and Levelt (2008), who report a correlation between the time-point children produce CVC sequences of a given place of articulation structure with the frequency of these structures in child-directed speech. In their longitudinal study, the high frequency of LC words produced by Dutch children in a certain stage is reflected in the high frequency of words with LC structures in their child-directed input. Although Japanese has an overall LC bias, plosives, the segment group that is among the first to be produced by infants, present a CL bias. Production data of Japanese infants and young children would therefore be a strong test of the hypothesis that assumes infants to start out with an LC bias in early production. If this were indeed found, a further step would require longitudinal data of Japanese children's productions, as learners of Japanese have to shift to a higher production rate of CL sequences eventually in order to get close to adult distributions.

## 8. Conclusions

Overall, our data support the notion that the productive LC bias is rooted in properties of the human articulatory system. However, perceptual preferences of these same sequences are influenced by distributional frequencies of the native language. There is no necessary perception–action link in the labial-coronal bias, and further language inven-

tories have to be studied in order to get a more complete picture of the pervasiveness of these biases.

## Acknowledgments

## References

Amano, S., & Kondo, T. (2000). *Nihongo-no Goi-tokusei* (vol. 7). Tokyo: Sansei-do (Lexical properties of Japanese).

Boersma, P., & Weenink, D. (2009). Praat: doing phonetics by computer (version 5.1.05).

Dupoux, E., Kakehi, K., Hirose, Y., Pallier, C., & Mehler, J. (1999). Epenthetic vowels in Japanese: A perceptual illusion? *Journal of Experimental Psychology: Human Perception and Performance, 25*, 1568–1578.

Fikkert, P., & Levelt, C. (2008). How does place fall into place? The lexicon and emergent constraints in children's developing phonological grammar. *Contrast in Phonology: Theory, Perception, Acquisition, 13*, 231–268.

Fougeron, C., & Smith, C. L. (1993). French. *Journal of the International Phonetic Association, 23*, 73–76.

Friederici, A. D., & Wessels, J. M. I. (1993). Phonotactic knowledge of word boundaries and its use in infant speech-perception. *Perception & Psychophysics, 54*(3), 287–295.

Gonzalez-Gomez, N., & Nazzi, T. (2012a). Phonological feature constraints on the acquisition of phonological dependencies. In *Proceedings of the 36th annual boston university conference on language development*. Somerville, MA: Cascadilla Press.

Gonzalez-Gomez, N., & Nazzi, T. (2012b). Acquisition of nonadjacent phonological dependencies in the native language during the first year of life. *Infancy.* http://dx.doi.org/10.1111/j.1532-7078.2011.00104.x.

Jusczyk, P. W., Friederici, A. D., Wessels, J. M. I., Svenkerud, V. Y., & Jusczyk, A. M. (1993). Infants sensitivity to the sound patterns of native language words. *Journal of Memory and Language, 32*(3), 402–420.

Jusczyk, P. W., Luce, P. A., & Charles-Luce, J. (1994). Infants sensitivity to phonotactic patterns in the native language. *Journal of Memory and Language, 33*(5), 630–645.

Lee, S., & Hasegawa, T. (2011). Bayesian phylogenetic analysis supports an agricultural origin of Japonic languages. *Proceedings of the Royal Society B.* http://dx.doi.org/10.1098/rspb.2011.0518.

Liberman, A. M., & Whalen, D. H. (2000). On the relation of speech to language. *Trends in Cognitive Sciences, 4*(5), 187–196.

Locke, J. L. (2000). Movement patterns in spoken language. *Science, 288*(5465), 449–451.

MacNeilage, P. F., & Davis, B. L. (2000). On the origin of internal structure of word forms. *Science, 288*(5465), 527–531.

MacNeilage, P. F., & Davis, B. L. (1998). The evolution of speech: The relation between ontogeny and phylogeny. In J. R. Hurford, C. Knight, & M. G. Studdert-Kennedy (Eds.), *The evolutionary emergence of language*. Cambridge: Cambridge University Press.

MacNeilage, P. F., Davis, B. L., Kinney, A., & Matyear, C. L. (1999). Origin of serial-output complexity in speech. *Psychological Science, 10*(5), 459–460.

MacNeilage, P. F., Davis, B. L., Kinney, A., & Matyear, C. L. (2000). The motor core of speech: A comparison of serial organization patterns in infants and languages. Invited submission to Special Millenium Issue. *Child Development, 71*(1), 153–163.

Maddieson, I. (1984). *Patterns of sounds*. Cambridge: Cambridge University Press.

Maekawa, K. (2003). Corpus of spontaneous Japanese: Its design and evaluation. In *IEEE workshop on spontaneous speech processing and recognition*, Tokyo.

McQueen, J. M. (1998). Segmentation of continuous speech using phonotactics. *Journal of Memory and Language, 39*(1), 21–46.

Mersad, K., & Nazzi, T. (2011). Transitional probabilities and positional frequency phonotactics in a hierarchical model of speech segmentation. *Memory & Cognition, 39*, 1085–1093.

Nazzi, T., Bertoncini, J., & Bijeljac-Babic, R. (2009). A perceptual equivalent of the labial–coronal effect in the first year of life. *Journal of the Acoustical Society of America, 126*(3), 1440–1446.

New, B., Pallier, C., Ferrand, L., & Matos, R. (2001). Une base de données lexicales du français contemporain sur internet: LEXIQUE. *L'Année Psychologique, 101*, 447–462. <http://www.lexique.org>.

Okada, H. (1991). Japanese. *Journal of the International Phonetic Association, 21*, 94–96.

Pena, M., Bonatti, L. L., Nespor, M., & Mehler, J. (2002). Signal-driven computations in speech processing. *Science, 298*(5593), 604–607.

Pitt, M. A., & Shoaf, L. (2002). Linking verbal transformations to their causes. *Journal of Experimental Psychology–Human Perception and Performance, 28*(1), 150–162.

Riney, T. J., Takagi, N., Ota, K., & Uchida, Y. (2007). The intermediate degree of VOT in Japanese initial voiceless stops. *Journal of Phonetics, 35*(3), 439–443.

Rochet-Capellan, A., & Schwartz, J. L. (2007). An articulatory basis for the labial-to-coronal effect: /pata/ Seems a more stable articulatory pattern than /tapa/. *Journal of the Acoustical Society of America, 121*(6), 3740–3754.

Saffran, J. R., Newport, E. L., & Aslin, R. N. (1996). Word segmentation: The role of distributional cues. *Journal of Memory and Language, 35*, 606–621.

Sato, M., Schwartz, J. L., Abry, C., Cathiard, M. A., & Loevenbruck, H. (2006). Multistable syllables as enacted percepts: A source of an asymmetric bias in the verbal transformation effect. *Perception & Psychophysics, 68*(3), 458–474.

Sato, M., Vallee, N., Schwartz, J. L., & Rousset, I. (2007). A perceptual correlate of the labial–coronal effect. *Journal of Speech Language and Hearing Research, 50*(6), 1466–1480.

Schwartz, J. -L., Basirat, A., Ménard, L., & Sato, M. (2012). The Perception-for-Action-Control Theory (PACT): A perceptuo-motor theory of speech perception. *Journal of Neurolinguistics, 25*(5), 336–354.

Vallée, N., Rousset, I., & Boë, L. J. (2001). Des lexiques aux syllabes des langues du monde. Typologies, tendances et organisations structurelles (From the lexicons to the syllables of the languages of the world. Typologies, tendencies and structural organizations.). *Linx, 45*, 37–50.

Vance, T. J. (1987). *An introduction to Japanese phonology*. Albany: State University of New York Press.

Vilain, A., Abry, C., Badin, P., & Brosda, S. (1999). From idiosyncratic pure frame to variegated babbling: Evidence from articulatory modelling. In *14th International Congress of Phonetic Sciences*, San Francisco.

Warren, R. M. (1961). Illusory changes of distinct speech upon repetition – The verbal transformation effect. *British Journal of Psychology, 52*(3), 249–258.

Warren, R. M., & Gregory, R. L. (1958). An auditory analogue of the visual reversible figure. *American Journal of Psychology, 71*, 612–613.

Weber, A., & Cutler, A. (2006). First-language phonotactics in second-language listening. *Journal of the Acoustical Society of America, 119*(1), 597–607.