No immersion, no instruction: Children's non-native vowel productions in a foreign language context.

Ellen Simon, Ronaldo Lima Jr & Ludovic De Cuypere

In: Proceedings of the International Symposium on Monolingual and Bilingual Speech (ISMBS), 2015, edited by Elena Babatsouli and David Ingram.

No immersion, no instruction:

Children's non-native vowel productions in a foreign language context

Ellen Simon[1], Ronaldo Lima Jr[2], Ludovic De Cuypere[1]

ellen.simon@UGent.be, ronaldojr@letras.ufc.br, ludovic.decuypere@UGent.be

[1]*Ghent University,* [2]*Federal University of Ceará*

Abstract. This study aims to map the native Dutch and non-native English vowels of Belgian children who have not been immersed and have not received any school-based instruction in English, but who are exposed to it through the media. A fairly large and recent body of research addresses second language perception and production by early learners either through immersion in an L2-speaking community or through classroom-based instruction. However, there is also a vastly expanding number of children who live in a monolingual community and yet are exposed to English as a Foreign Language (EFL) from an early age through various media. This study addresses the question to what extent children acquire the English vowel system in such a context: is this type of exposure sufficient for them to create new phonetic vowel categories? Twenty-four Dutch-speaking children, aged between 9 and 12, participated in the study. They were all living in Belgium, and came from different dialect regions. None of them had received English instruction in school, but all of them reported having at least some sporadic contact with English, for instance through television programmes or computer games. They all performed two Dutch picture-matching tasks, an English repetition task, and an English picture-naming task. The auditory stimuli were monosyllabic Dutch and English words containing each of 12 Dutch and 11 English monophthongs. The vowel formants were analysed in Praat (Boersma & Weenink, 2014) by comparing the LPC (Linear Predictive Coding) analysis to the FFT (Fast Fourier Transform) spectrum. Lobanov-normalized vowel plots present the organization of these children's entire Dutch and English vowel spaces. The results focus on the English vowel contrasts /ɛ-æ/ and /ʊ-u/, as Dutch lacks these contrasts and has only one vowel in these areas of the vowel space (/ɛ/ and /u/, respectively). The children produced a contrast between English /ɛ/ and /æ/ in the repetition task, but not in the picture-naming task. English /ɛ/, but not /æ/ was considerably different from the closest Dutch vowel /ɛ/. The children's English /ʊ/ and /u/ differed in terms of height (F1) and anteriority (F2), both in the repetition and the picture-naming task. The closest Dutch vowel, represented as /u/, did not differ from English /u/, and differed from /ʊ/ only in terms of height. The results suggest that 9-12-year-old Flemish children are at the beginning of creating new contrasts for non-native English vowels. This means that media-induced Second Language Acquisition should not be underestimated: even in contexts of L2 acquisition exclusively through media exposure children learn to produce contrasts between L2 vowels which do not exist in their L1.

Keywords. Child second language phonology, vowels, production, acoustics, Dutch, English.

Introduction and aims

This study aims to map the native (L1) and non-native (L2) vowels of children who have not yet received any school-based instruction in the L2, but who have been exposed to it in a non-immersion context. Studies on L2 phonological acquisition have typically focused on immersion contexts, often examining language acquisition by immigrants. In these contexts, once L2 acquisition starts, it is typically with intense exposure. The results of these studies (e.g. Tsukada, Birdsong, Bialystok, Mack, Sung & Flege, 2005; Gildersleeve-Neuman, Peña, Davis & Kester, 2009; Darcy & Krüger, 2012) show that the L1 is generally still permeable in childhood, and that the children's L2 productions differ not only from those in their L1, but also from those of age-matched L1 children. Another set of studies on child L2 acquisition have focused on the effect of instruction on child L2 phonological acquisition, mostly examining the effect of age of onset of instruction on the attained proficiency level. The Barcelona Age Factor project (Muñoz, 2006), conducted longitudinally between 1996 and 2002, compared pupils for whom English instruction started at age 11 to pupils who started getting English instruction at age 8. Muñoz' conclusion of the project as a whole is that no group of learners performed even close to the native speakers that composed the control group. Late starters (age 11) performed better than early starters (age 8) at all phases of data collection, but the older learners' advantage decreased in the later collections. Conclusions of studies within the project focusing on perception and production (Fulana, 2006) and oral fluency (Álvarez, 2006; Mora, 2006) reached the same conclusion.

These studies suggest that, in contexts of maximal input, either through immersion or intensive instruction or training, children's L2 speech is influenced by their L1 and differs from that of age-matched L1 speakers. The question we address in this paper is what child L2 speech looks like in contexts of minimal input, i.e. in the absence of immersion or formal instruction. Such contexts are actually common: in many European countries, including Belgium, children are exposed to English through various media, such as computer games, television programmes and the radio, before they get English classes in school.

In this study we examine to what extent 9-12-year-old Dutch-speaking children living in Flanders have acquired the spectral quality of L2 English vowel sounds as the result of exposure to English through various media. Since children are exposed to multiple varieties of English (as is typical for English as a Foreign Language contexts, see Bohn & Bundgaard-Nielsen, 2007), the children's L2 vowels will not be compared to those of a control group of L1 speakers. Rather, we examine the internal organization of the children's L1 and L2 vowels spaces. In this paper, we will zoom in on two L2 vowel contrasts which do not occur in the L1, and address the following questions: (1) Do the children produce a contrast between the L2 English vowels in these pairs?, and (2) Do these productions differ from the closest L1 Dutch vowel?

Method

Participants

Twenty-four Dutch-speaking children, living in Flanders, Belgium, participated in Dutch and English production tasks. The mean age of the participants (9 girls, 15 boys) at the time of testing was 10;6 years (range: 9;10 to 12;2). Data were collected in three schools in different towns in Flanders, Ghent (n = 9), Erembodegem (n = 6) and Mol (n = 9), in order to examine potential effects of L1 regional variation. None of the children had received any formal L2 English instruction in school or made extended trips to English-speaking countries and no children reported having contact with native English speakers. However, all children in Belgium are exposed to English through the media and popular culture (music channels, English-spoken cartoon channels, computer games, English pop music, etc.), so that by the age of 9, they have a basic knowledge of English.

Tasks and procedure

All children performed a Dutch picture-matching task, an English repetition task, and an English picture-naming task. In the Dutch picture-matching task, they were asked to match pictures while producing sentence of the form 'X belongs to Y', in which either X or Y was a target word (e.g. 'The *cheese* belongs to the mouse' - 'De *kaas* hoort bij de muis').

In the English repetition task, children saw pictures on a computer screen and heard the corresponding words over Bose™ headphones. They were instructed to repeat the words. The audio recordings, produced by a male and a female speaker of British English, were extracted from the online version of the Cambridge Advanced Learner's Dictionary (Upper intermediate – advanced) (Cambridge: Cambridge University Press, third edition, http://dictionary.cambridge.org). The English picture-naming task aimed at eliciting spontaneously produced words as opposed to repeated words. Children were shown six cards with four pictures on each and were asked to name the objects for which they knew the names in English.

Experimental set-up

The children were individually tested in a quiet room in their school, with no other person present besides the experimenter. All instructions were provided orally in Dutch. The recordings were made with a Sony clip microphone (ECMCS10), connected to a pocket-size Marantz Professional solid state recorder (PMD620). The recordings were made in Mono, with a sampling rate of 44.1KHz. All tasks were performed in one session, and always in the order in which they are presented above.

Stimuli

All visual stimuli were black or coloured line drawings, taken from the web. The auditory stimuli were monosyllabic Dutch and English basic vocabulary words. Monosyllabic words with each of the 12 Dutch (/ɛ/, /ɪ/, /u/, /ɑ/, /i/, /ɔ/, /a/, /ʏ/, /o/, /ø/, /e/, /y/) and 11 English monophthongs (/ɛ/, /ɪ/, /u/, /ɑ/, /i/, /ɔ/, /æ/, /ʊ/, /ɜ/, /ʌ/, /ɒ/) were selected, excluding schwa. Since the children's vocabulary in English was very limited, the consonantal context of the words could not be controlled for. All target words were high-frequency English words likely to be known by the majority of the children (mean log frequencies: picture-naming task: 9.954, SD 1.19; repetition task: 9.93, SD 1.27; frequencies from Balota et al., 2007).

Analysis

The spectral analysis is based on measurements of the first and second formants. After the vowels were segmented in PRAAT (Boersma & Weenink, 2011), formant values predicted by LPC (Linear Predictive Coding) were manually checked against the FFT power spectrum (obtained by the calculation of the Fast Fourier Transform algorithm) of the central, most stable part of each vowel. This manual checking allowed for adjustments to be made in the ceiling frequency and/or the order of the LPC whenever necessary, which is essential when working with children, whose ceiling frequencies may vary considerably from one to another due to their still developing vocal tracts and typical high F0 values. A PRAAT script (Arantes, 2010) was used to visualize the LPC predictions against the FFT spectrum, and to change the parameters of analysis when necessary, and another script (Arantes, 2011) was used to later export all resulting F1 and F2 values to a spreadsheet. After extraction, F1 and F2 values were Lobanov normalized (Lobanov, 1971) and the output values were rescaled to Hertz, using the 'vowels' package (Kendall & Thomas, 2009) for R software (R Core Team, 2012). On the basis of visual inspection of the scatterplots (see Figures 1 and 3 in section 4), we identified 60 vowel productions with extreme values. After a close, manual examination of these 60 vowels, 49 observations were removed because background noise or extreme lengthening or whispering made the measurement unreliable. Thus, extreme values were deleted for technical reasons only, not because of their distance from the bivariate means. The normalized data were then used to create F1xF2 plots and to conduct joint multivariate tests.

In total, 793 Dutch and 1303 English vowels were retained in the analysis, leading to a total of 2096 vowels. For this paper, we focus on the analysis of two English vowel contrasts, which do not occur in Dutch. In these two pairs, Dutch has just one vowel in the area of the vowel space where English has two (see Table 1), and both pairs are hence predicted to be problematic for native speakers of Dutch.

Table 1. Three English vowel contrasts and the spectrally closest Dutch vowel.

	English pairs	Closest Dutch vowel
1.	/ɛ-æ/ (‘DRESS’-‘TRAP’)	/ɛ/ (‘MES’)
2.	/ʊ-u/ (‘FOOT’-‘GOOSE’)	/u/ (‘HOEK’)

Results

DRESS – TRAP vs. MES

Figure 1 presents a scatterplot of all productions of the English vowels /ɛ/ (‘DRESS’) and /æ/ (‘TRAP’) (left) as well as the closest Dutch vowel /ɛ/ (‘MES’) (right). All scatter plots are created with McCloy’s (2015) PhonR package in R. The leftmost panel includes results of the picture-naming as well as repetition task. The rightmost panel includes only the English and Dutch picture-naming/matching task, since no repetition task was conducted in Dutch.

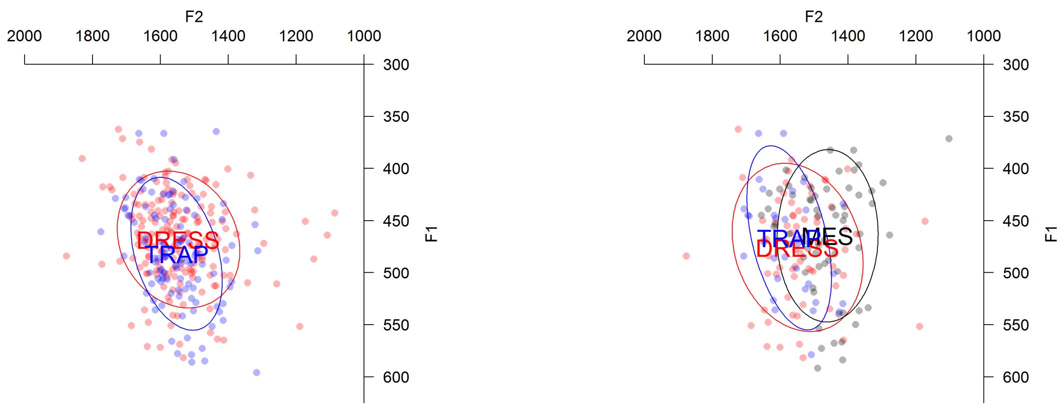

Figure 1. Scatterplot of English DRESS and TRAP (left) (spontaneous and repetition tasks), and comparison with Dutch MES (right) (spontaneous task only).

The scatter plots suggest a difference between DRESS and TRAP on F1 and a difference between MES and TRAP/DRESS on F2. The results of a joint multivariate test on the bivariate means for English DRESS and TRAP, controlling for TASK and REGION, show a significant effect of TARGET VOWEL in interaction with TASK (repetition vs. picture-naming/matching; Type II MANOVA: Hotelling-Lawley test, $P = 0.02$). (All statistical analyses were performed in R).

A post-hoc linear regression analysis on both formants separately indicates that TARGET VOWEL was significant in interaction with TASK for F1 ($P < 0.01$). The interaction plot in Figure 2 shows that F1 for TRAP is much higher than for DRESS in the repetition task, which is expected in English, but the reverse pattern can be observed in the picture-naming/matching task. While the 95% confidence intervals (the red bars) do not overlap in the repetition task, they do overlap in the spontaneous task, meaning that in the picture-naming task (referred to as the ‘spontaneous’ task) there is no evidence that a contrast is being made.

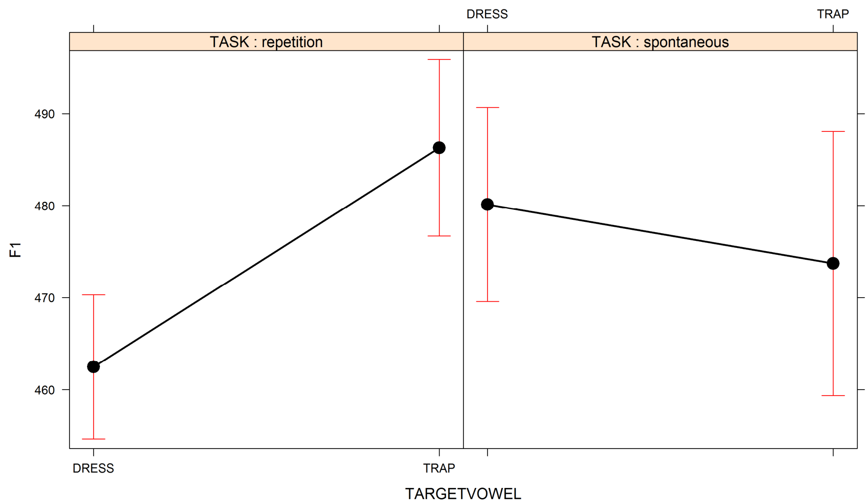

Figure 2. Interaction plot for TASK and TARGET VOWEL for F1.

No difference between the target vowels was found for F2, which is in line with what the scatterplot in Figure 1 shows.

A multivariate comparison of DRESS and TRAP with the closest Dutch vowel, MES, again revealed a significant effect of TARGET VOWEL (Type II MANOVA test: Pillai test, p < 0.001). The post-hoc linear regression model showed that Dutch MES was significantly different from English DRESS in terms of F2 (P < 0.001)and F1 for the REGION Erembodegem. The difference with TRAP was not significant, neither in F1 nor F2.

FOOT-GOOSE vs. HOEK

Figure 3 presents a scatterplot of all productions of the English vowels /ʊ/ (‘FOOT) and /u/ (‘GOOSE’) (left) as well as the closest Dutch vowel /u/ (‘HOEK’) (right).

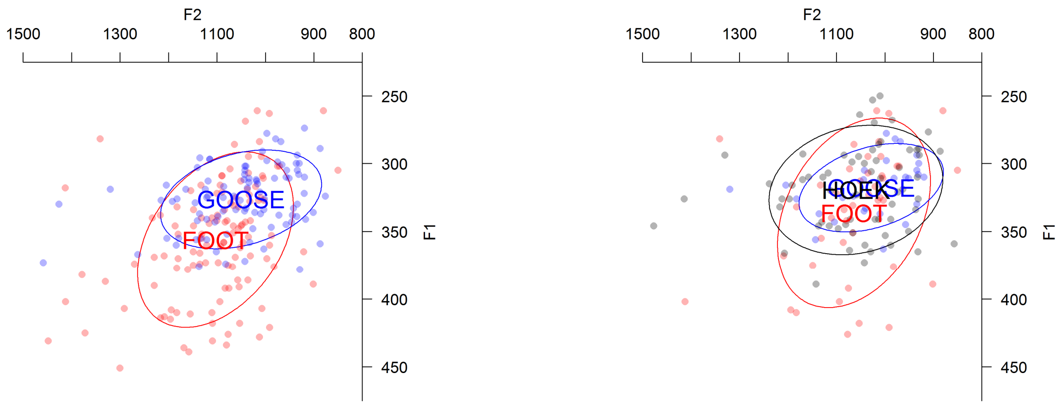

Figure 3. Scatterplot of English FOOT and GOOSE (left) (spontaneous and repetition task), and comparison with Dutch HOEK (right) (spontaneous task only).

As for the DRESS-TRAP contrast, a joint multivariate test on English FOOT and GOOSE productions revealed a highly significant effect of TARGET VOWEL, controlling for REGION and TASK (Type II MANOVA, Hotelling-Lawley test: P < 0.001). A post-hoc linear regression analysis confirmed that the two vowels differed significantly both in F1 and F2 (P < 0.001), again controlling for REGION and TASK.

A comparison with the closest Dutch vowel, HOEK, showed no evidence of a multivariate difference between the three vowels means (Type II MANOVA, Pillai test: P = .054). However, a post-hoc linear regression analysis revealed that Dutch HOEK was different from English FOOT in terms of F1 ($p = 0.02$), but not in terms of F2. No difference between HOEK and GOOSE was found in either F1 or F2.

Discussion and conclusions

This study addressed the question whether Dutch-speaking children living in Flanders learn to create new categories for English vowels before they have received English instruction in school. In other words, is sheer exposure to English-spoken media sufficient for children to develop new L2 vowel categories, and to what extent do these vowel categories differ from the spectrally closest L1 Dutch vowels? For this paper, we zoomed in on two English vowel contrasts which do not occur in Dutch, namely /ɛ-æ/ and /ʊ-u/. Even though the DRESS-TRAP contrast is known to be difficult for native speakers of Dutch, both in perception (Broersma, 2005; Escudero, Simon & Holgerer, 2012) and in production (Simon & D'Hulster, 2012), children produced these English vowels significantly different, both in terms of F1 and F2, but only in a repetition task. We found no evidence for a contrast between DRESS and TRAP in a picture-naming task, in which children had to retrieve their phonological representations of the L2 vowels. A comparison with the closest Dutch vowel, MES, conventionally represented by the phonetic symbol /ɛ/, showed that the children produced this Dutch vowel differently from English /ɛ/, both in terms of height and anteriority, but not different from English /æ/. With respect to the FOOT-GOOSE contrast, the results again showed that children produced a contrast between these vowels, both in terms of height and anteriority, and this time they did so both in the repetition and the picture-naming task. The closest Dutch vowel, HOEK, represented as /u/, did not differ from English GOOSE, and differed from FOOT only in terms of height. In other words, even though the children's Dutch vowel is highly similar to both English vowels, the children managed to produce a contrast between these two L2 vowels.

To conclude, the results suggest that 9-12-year-old Flemish children are at the beginning of creating new contrasts for non-native English vowels. This means that media-induced Second Language Acquisition should not be underestimated: even in contexts of L2 acquisition exclusively through media exposure ('no immersion - no instruction'), children learn to produce contrasts between L2 vowels which do not exist in their L1. The results are interesting in light of the relation between perception and production. A previous perception study with the same group of Flemish children (Simon, Sjerps & Fikkert, 2012), based on mispronunciation detection tasks, showed that the children's perception of L2 English vowels was strongly influenced by their L1, but that the beginning of development of new categories could be detected. However, while the children are exposed to English-spoken media from an early age onwards, and get a considerable amount of L2 receptive input, they hardly ever produce the L2. Interviews with the child participants revealed that production of English was restricted to singing along with pop songs and the use of occasional English phrases with friends. Yet, despite this lack of productive practise, the children are at the beginning of creating new categories in their production, on the basis of their receptive input.

In addition, the results may have a pedagogical impact: children who are not immersed in the L2 and have not even had English classes in school yet, have an L2 vowel space which is different from their L1 vowel space, which is something that teachers in the first years of English language instruction in school may want to take into account when developing their teaching materials.

References

Alvarez, E. (2006). Rate and route of acquisition in EFL narrative development at different ages. In: C. Munoz (ed.). *Age and the Rate of Foreign Language Learning* (pp. 127-155). Clevedon, Tonawanda NY & Ontario: Multilingual Matters Ltd.

Arantes, P. (2010). Formants.praat, v. 0.9 beta.

Arantes, P. (2011). Collect formants.praat, v. 0.11 alpha.

Balota, D.A., Yap, M.J., Cortese, M.J., Hutchison, K.A., Kessler, B., Loftis, B., Neely, J.H., Nelson, D.L., Simpson, G.B., & Treiman, R. (2007). The English Lexicon Project. *Behavior Research Methods*, *39*, 445-459.

Boersma, P. & Weenink, D. (2011). Praat: Doing Phonetics by Computer. [Computer Programme], http://www.praat.org.

Bohn, O.-S. & R.L. Bundgaard-Nielsen. (2007). Second language speech learning with diverse input. In: T. Piske & M. Young-Scholten (eds.), *Input Matters in SLA* (pp. 207-218), Clevedon: Multilingual Matters.

Broersma, M. (2005). Perception of familiar contrasts in unfamiliar positions. *Journal of the Acoustical Society of America, 117*, 3890-3901.

Darcy, I. & F. Krüger. (2012). Vowel perception and production in Turkish children acquiring L2 German, *Journal of Phonetics, 40*, 568-581.

De Jans, K. (2013). *The lexical knowledge of English of secondary school children in Flanders*. Unpublished Ma thesis, Ghent University.

Escudero, P., Simon, E. & Mitterer, H. (2012). The perception of English front vowels by North Holland and Flemish listeners: Acoustic similarity predicts and explains cross-linguistic and L2 perception, *Journal of Phonetics, 40*, 280-288.

Fulana, N. (2006) The development of English (FL) perception and production skills: starting age and exposure effects. In: C. Munoz (ed.). *Age and the Rate of Foreign Language Learning* (pp. 41-64). Clevedon, Tonawanda NY & Ontario: Multilingual Matters Ltd.

Gildersleeve-Neuman, C.E., Peña, E.D., Davis, B. & Kester, E. (2009). Effects of L1 during early acquisition of L2: Speech changes in Spanish at first English contact, *Bilingualism: Language and Cognition, 12 (2),* 259–272.

Kendall, T. & Thomas, E. R. (2010). Vowels: Vowel manipulation, normalization, and plotting in R. R package, v. 1.1, [Software available online: http://ncslaap.lib.ncsu.edu/tools/norm/].

Lobanov, B.M. (1971). Classification of Russian vowels spoken by different listeners. *Journal of the Acoustical Society of America, 49*, 606-08.

McCloy, D. (2015). PhonR: tools for phoneticians and phonologists. R package version 1.0-3.

Mora, J. C. (2006). Age effects on oral fluency development. In: C. Munoz (ed.). *Age and the Rate of Foreign Language Learning* (pp. 65-88). Clevedon, Tonawanda NY & Ontario: Multilingual Matters Ltd.

Munoz, C. (2006). (Ed.) *Age and the Rate of Foreign Language Learning*. Clevedon, Tonawanda NY & Ontario: Multilingual Matters Ltd.

R Core Team. (2014) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Online: http://www.R-project.org/.

Simon, E., Sjerps, M. & Fikkert, P. (2013). Phonological representations in children's native and non-native lexicon, *Bilingualism: Language and Cognition, 17.1*, 3-21.

Simon, E. & D'Hulster, T. (2012). The effect of experience on the acquisition of a non-native vowel contrast, *Language Sciences, 34,* 269-283.

Tsukada, K., Birdsong, D., Bialystok, E., Mack, M., Sung, H. & Flege, J. (2005). A developmental study of English vowel production and perception by native Korean adults and children, *Journal of Phonetics, 33*, 263–290.