



HAL
open science

A new interactive graphical approach to visualize children phoneme's acquisition over time

Andrea Briglia, Jérémie Sauvage

► To cite this version:

Andrea Briglia, Jérémie Sauvage. A new interactive graphical approach to visualize children phoneme's acquisition over time. *Phonetics and Phonology in Europe 2021*, Universitat de Barcelona, Jun 2021, Barcelona, Spain. <https://pape2021.upf.edu/program-4/>. hal-03565898

HAL Id: hal-03565898

<https://hal.science/hal-03565898>

Submitted on 16 May 2024

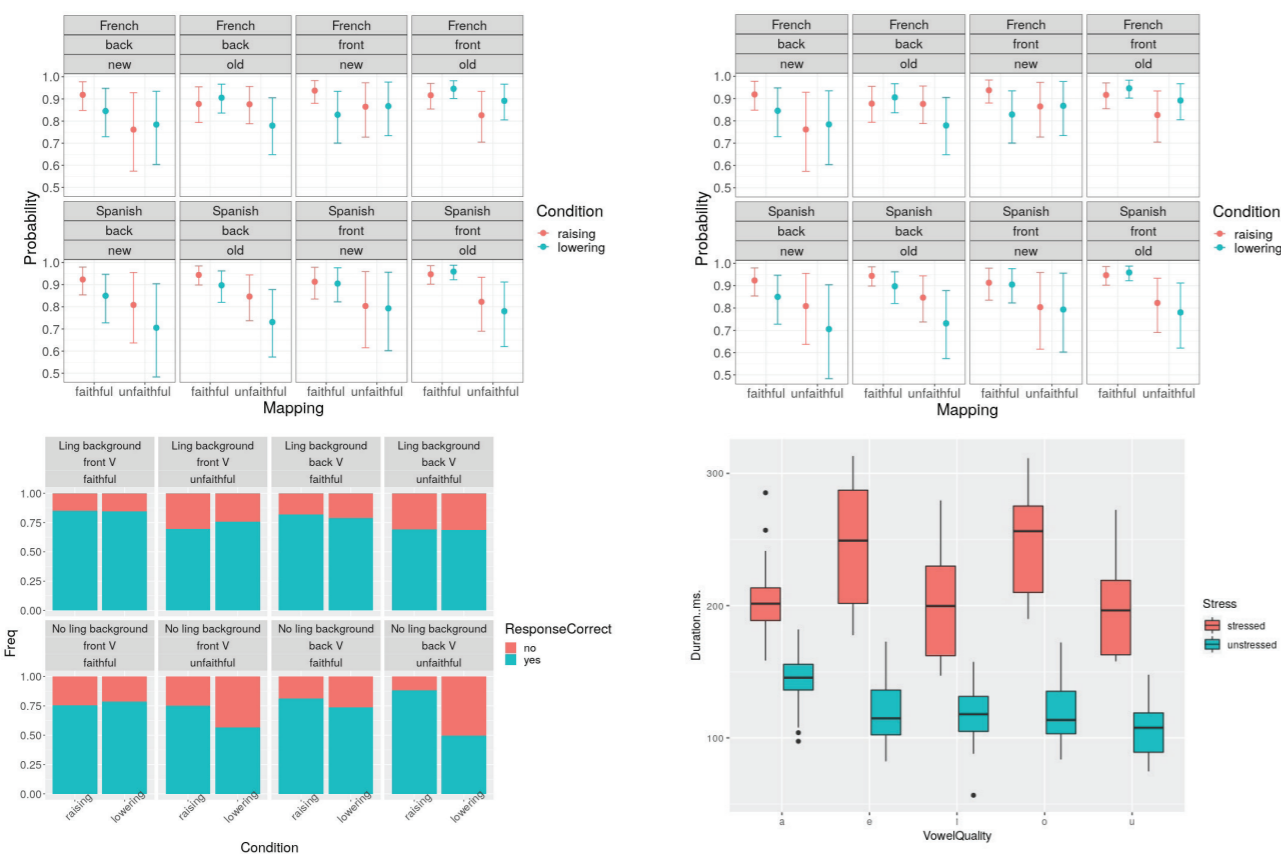
HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

.95). Moreover, the (non-significant) advantage for raising appears to be largely limited to back vowels. We advance two tentative explanations for these inconclusive results. First, we noted that the asymmetry favoring raising appeared to be stronger for participants who declared having no educational background in linguistics (n=28), in particular for back vowels, as shown in Figure 2a. We speculate that linguistically informed participants relied on metalinguistic awareness more than naive participants when performing the task and were therefore less sensitive to phonetic detail. To control for this confound, we plan to rerun the same study and include only linguistically naive participants. Concerning the stronger advantage for raising with back vowels, we hypothesize that it is due to a phonetic asymmetry in our stimuli. As shown in Figure 2b, high back vowels were generally shorter than their mid counterparts in unstressed syllables but high front vowels were not shorter than their mid counterparts in unstressed syllables. If high vowels are better candidates for unstressed vowels as long as they are clearly shorter than the corresponding mid vowels, then it makes sense that raising would be easier to learn for back vowels (where raising correlates with shortening) than for front vowels (where raising does not correlate with shortening).



- [1] Gordon, M. K. 2016. *Phonological typology*. Oxford: Oxford University Press.
- [2] Crosswhite, K. M. 2004. Vowel reduction. In Hayes, B., Kirchner, R. & Steriade, D. (Eds), *Phonetically based phonology*. Cambridge: Cambridge University Press, 191-231.
- [3] Lehiste, Ilse. 1970. *Suprasegmentals*. Cambridge, MA: The MIT Press.
- [4] Lindblom, B. 1963. Spectrographic study of vowel reduction. *Journal of the Acoustical Society of America* 35, 1773-1781.
- [5] Martin, A. & Peperkamp, S. 2020. Phonetically natural rules benefit from a learning bias: A re-examination of vowel harmony and disharmony. *Phonology* 37, 65-90.
- [6] Martin, A. & White, J. 2021. Vowel harmony and disharmony are not equivalent in learning. *Linguistic Inquiry* 52, 227-239.
- [7] Bürkner, P.-C. 2017. brms: An R package for bayesian multilevel models using stan. *Journal of Statistical Software* 80, 1-28.

A new interactive graphical approach to visualise children phonemes' acquisition over time

Briglia A., Sauvage J.
LHUMAIN, Université « Paul-Valéry » Montpellier (France)

Data. CoLaJE is a French open access database (available in CHILDES) containing longitudinal recordings of *in vivo* child spoken language from seven children that have been regularly filmed one hour every month from their first year of life until 5. Each record has been transcribed in CHAT, for two children an additional transcription in orthographic norm and IPA is provided. Data is coded in two tier : "pho" represents what the child says and "mod" represents what the child should have said according to the adult's norm. This sampling scheme meets child language representativeness standards (Stahl & Tomasello, 2004). Data is turned into a machine-readable format by converting it into .txt to make it suitable for Python.

Method. We wrote a set of scripts to extract every occurrence of each of the 36 phonemes specific to French. It would have been possible to code Child-Directed Speech too, as everything adults say has been transcribed in these *corpora*, but we decided to focus on child production only, even considering that the data amount was already huge. We ordered the 36 phonemes into a hierarchical list according to their articulatory effort (Sauvage, 2015). Aim was to obtain a global overview of the phonemes' acquisition over age and, at the same time, having the possibility to focus on the specific development of any given phoneme in relation to every other one : to do so, we chose to adopt the Multiresolution Streamgraph (Cuenca et al., 2018). This graph has been conceived to overcome readability and scalability issues in the visualisation of multiple time series : it allows a navigation through the different hierarchical levels providing a way to aggregate and disaggregate its structure, as well as to filter and highlight a specific element. It is possible to know how many occurrences of any phoneme has been pronounced by any of the seven children in any of the recording sessions. Results are coherent to current literature on consonants acquisition (McLeod et al., 2018) in the sense that they approximately fit to the state-of-the-art research on the order of acquisition of consonant and vowels. A comparison has been made on the children of this dataset, allowing us to establish the amount of differences and similarities between them by overlapping graphs in a synoptic way. Results obtained were coherent to previous studies on French (Yamaguchi, 2012). The advantage of this graph is its versatility, the main limit is that it does not consider coarticulatory differences. As a number of phonemes differ in their pronunciation according to the place they occupy (onset, nucleus, coda), the visualisation is somehow coarse-grained. Tackling the problem of coarticulatory differences with programming languages is currently one of the most difficult task to solve in NLP.

Results and discussion. Among current theories on phonological acquisition, in particular OT (Prince, 1993), we thought that Clements' « Theory of phonological features » was the closest to be complementary to the study of frequency effects in language acquisition, especially for its quantitative groundedness. We tested whether the occurrences in this graph fit to the three pillars of this theory : according to the « markedness avoidance » principle, we found that marked phonemes occur most of the times less than unmarked ones (e.g « p » occurs more times than « b », « t » occurs more times than « d », « k » more times than « g » and so on). By focusing on the same group of consonants specified by the voicing contrast, we found how the « feature economy » principle - according to which features present in one

segment tend to be used to define others - is coherent to what it is possible to observe from the graphs : as this feature specifies more phonemes than any other one in French, those features appear more frequently and - *a fortiori* - children prioritize their acquisition. The same holds for principle of feature hierarchy : consonants whose defining traits are placed at the root are learnt most of the time before consonants whose defining traits are placed in subsequent branches (Yamaguchi, p130). These preliminary results are part of a Phd project : we are currently working on the analyses of the same *corpora* through features instead of phonemes. By visualising the hierarchy of features in the Multistream, we suppose to improve current knowledge on the path of phonological variations (*i.e* the sequence of temporary achieved stages toward the adult norm). To give an example, « tʁaktœv » is never pronounced « dʁaktœv » but instead « kʁaktœv » as « k » is voiceless as « t ». In this way we would try to improve knowledge about the relation between feature hierarchy and acquisition. As far as we know, the visualisation technique we propose is innovative and could be used as a tool to help researchers to answer questions about the L1 acquisition process in any language.

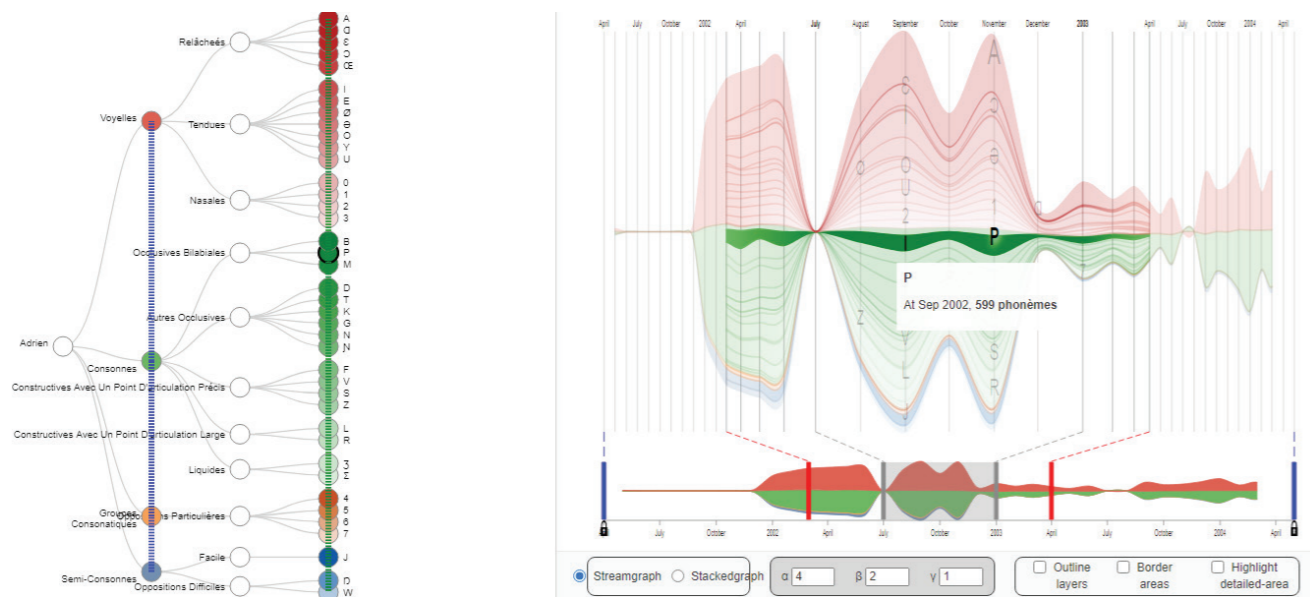


Figure 1. Multiresolution Streamgraph obtained from Adrien dataset <https://marine27.github.io/TER/index.html>

References

- [1] Morgenstern A.; Parisse C. (2012), « *The Paris Corpus* ». French language studies 22. 7-12. Cambridge University Press. Special Issue.
- [2] Tomasello, M. and Stahl, D. (2004). « *Sampling children's spontaneous speech : How much is enough ?* ». Journal of Child Language, 31 :101-121
- [3] Sauvage J. (2015). « *L'acquisition du langage. Un système complexe*. L'Harmattan- Louvain-la-neuve.
- [4] Cuenca E., Sallaberry A., Wang Y., Poncelet P. (2018). « *MultiStream : A Multiresolution Streamgraph Approach to explore Hierarchical Time Series* ». IEEE Transactions on visualization and computer graphics, vol.24, no. 12
- [5] McLeod S.; Crowe K. (2018). "Children's Consonant Acquisition in 27 Languages: A Cross-linguistic Review". American Journal of Speech-Language Pathology. P 1-26
- [6] Clements, G. N. & R. Ridouane (2011). "Where do phonological features come from? Cognitive, physical and developmental bases of distinctive speech categories". John Benjamins. Amsterdam

DAY 1 - Poster Session C Prosody: Information Structure

Prosodic expression of information status in Italian

Simona Sbranna, Caterina Ventura, Aviad Albert, Martine Grice
Universität zu Köln

Keywords: prosodic marking, information status, periogram, Italian.

Introduction. Previous research investigating prosodic marking of information status claims that Italian tends to resist deaccentuation of given elements [1, 2, 3]. In particular, Italian reportedly accents post-focal given information within noun phrases (NPs), making it difficult to reconstruct the information status of items from the acoustic signal [4]. However, so far there have only been categorical descriptions of accent patterns, risking the loss of important information about the continuous phonetic parameters that contribute to pitch accent categorisation. Since subtle modulations of continuous prosodic parameters may play a role in differentiating information structures, this paper uses a novel approach to explore how speakers modulate prosodic cues for this purpose. The variety of Italian investigated is that spoken in Naples.

Method. We designed a board game to elicit utterance-final NPs with disyllabic nouns and adjectives evoking different information structures – given-new (GN), new-new (NN), new-given (NG) – in statements. A total of 774 analysable items were obtained from 54 speakers. We investigated continuous periodic energy and F0 trajectories to achieve enriched visual representations of pitch contours, *i.e.* periograms [5], and to quantify F0 contour shapes in temporal and spectral dimensions via *synchrony* and *scaling* [6, 7]. Synchrony is the distance between the tonal centre of gravity [8] and the centre of periodic mass within syllabic intervals. It reflects the perception of a rising or falling F0 shape within a syllable. Scaling reflects the difference between F0 values at the centre of periodic mass across two consecutive syllables. To validate the production data, three phoneticians, native speakers of the Neapolitan variety, listened to the analysed items in a context-matching perception study, indicating whether items were perceived as GN, NN, or NG.

Results. Example periograms and periodic energy curves for the three conditions [Fig. 1] reveal two patterns (GN and NN vs. NG) reflected in the averaged values for synchrony and scaling in the entire corpus [Fig. 2.a]. GN and NN items have a similar contour with a late F0 fall on the second syllable, while NG items have an earlier F0 fall, on the first syllable. In all conditions the F0 contour on the adjective (third and fourth syllables) is relatively flat, starting with a slight fall in the GN and NN conditions. These patterns are reflected in the synchrony and scaling values [Fig. 2.a]. In GN and NN the positive synchrony value on the first syllable signals a rising movement, and the negative value on the second syllable indicates a mostly falling movement. In addition, the positive scaling value on the second syllable confirms a rise from the first to the second syllable. By contrast, in NG the F0 fall already starts on the first syllable, reflected in the negative scaling value on the second syllable, suggesting a falling movement across the first and second syllables. In both GN and NN the F0 reaches its lowest point on the third syllable and remains stable across the fourth syllable, as indicated by the scaling values approaching zero. Results from the context-matching study [Tab. 1] show that GN and NN items are frequently confused, whereas agreement was consistently more robust for NG items, reflecting the patterns in synchrony and scaling. Fig. 2.b displays the values of synchrony and scaling for the items that were correctly matched by all three listeners. Here