

# Accommodation temporelle chez l'enfant dans une tâche de parole alternée

Céline Hidalgo<sup>1,2</sup> Simone Falk<sup>1,3</sup> Daniele Schön<sup>2</sup>

(1) Laboratoire Parole et Langage, UMR 7309, CNRS, Aix-en-Provence, France

(2) Institut de Neurosciences des Systèmes UMR 1106, INSERM, Marseille, France

(3) Institut de Philologie allemande, Ludwig-Maximilians-Universitaet, Munich, Allemagne

celine.hidalgo@univ.etu-amu.fr, Simone.Falk@germanistik.uni-muenchen.de,  
daniele.schon@univ-amu.fr

## RESUME

---

L'accommodation temporelle entre deux interlocuteurs est un phénomène qui émerge lors d'une interaction et qui jouerait un rôle important dans la fluidité des échanges. Cette étude examine cette capacité temporelle chez l'enfant âgé de 5 à 6 ans grâce au développement d'une nouvelle tâche de dénomination en alternance avec un partenaire virtuel. Les variables temporelles analysées sont le tempo de l'alternance (*lent* versus *rapide*) et la rythmicité des mots échangés (*constante* versus *aléatoire*). Les enfants sont plus précis dans la condition de tempo rapide et plus réguliers lorsque la rythmicité des listes de mots est maintenue constante. Ces résultats montrent 1) que la dénomination en alternance est un paradigme permettant de mesurer les capacités d'accommodation temporelle des enfants et que 2) dès 5 ans, les enfants peuvent ajuster leur parole à celle d'un agent. Ces données constituent une base pour mesurer les capacités linguistiques d'accommodation temporelle chez des populations cliniques.

## ABSTRACT

---

### **Children's temporal accommodation in an alternated naming task.**

Temporal accommodation between interlocutors is a phenomenon taking place in language interactions that may play an important role in improving communication. This study examines temporal accommodation skills of children aged from 5 to 6 years using a new paradigm. In a picture naming task in alternation with a virtual agent, we analyzed children's temporal accommodation according to two temporal parameters : speed of alternation - *fast* or *slow* - and the number of syllables of the words - *match* or *mismatch* - pronounced by the agent and children. Children were more accurate in the fast condition and more regular in the match condition. These results show that 1) naming in alternation is an effective task to measure temporal accommodation of children and that 2) children from age 5 adjust their speech to the timing of an agent. This study constitutes a baseline to assess language interaction abilities in clinical populations.

**MOTS-CLES** : accommodation interpersonnelle, enfants, prédiction temporelle, tour de parole, interaction, agent.

**KEYWORDS** : interpersonal accommodation, children, timing prediction, turn-taking, interaction, agent.

# 1 Introduction

La conversation est une situation co-construite par les interlocuteurs qui évolue au cours du temps. La dynamique de cette co-construction, est sous-tendue par un phénomène ténu mais néanmoins essentiel : la convergence. Des études ont en effet montré qu'au cours d'une interaction, les locuteurs ont une tendance à imiter certaines caractéristiques verbales- (phonétiques : Pardo, 2006 ; syntaxiques : Branigan et al., 2000)-et paraverbales (Krivokapić, 2013) de leur interlocuteur ou encore à s'aligner conceptuellement avec lui (Garrod & Pickering, 2004). Ces procédés nommés *convergence*, *alignement*, *accommodation*, *synchronisation*, permettraient d'une manière générale de favoriser une bonne compréhension réciproque. Plus spécifiquement, ces phénomènes appliqués au niveau temporel permettraient de fluidifier la dynamique conversationnelle. En effet, en mimant certaines caractéristiques temporelles de la parole de son interlocuteur, l'auditeur augmenterait sa capacité à prédire la fin des tours de parole. Ces prédictions temporelles lui permettraient ainsi d'anticiper la programmation de sa propre parole afin qu'elle soit produite sans trop de délai ni trop de chevauchement lors de son tour de parole.

Selon Wilson & Wilson (2005), ce mécanisme de projection lors de la prise des tours entre les interlocuteurs se réalise au niveau cérébral grâce au couplage de phase (phase-locking) entre la fréquence d'apparition des syllabes du locuteur et l'activité oscillatoire spontanée de ses neurones. Ainsi, le système moteur du locuteur et le système auditif de l'auditeur entreraient en couplage de phase à partir de la fréquence du rythme de parole (débit ou nombre de syllabes par seconde) qui est déterminé de manière dynamique (donc variable) au cours de l'interaction par les mouvements oscillatoires de la mandibule des interlocuteurs (Scott et al., 2009). Ce phénomène de « phase-locking » neuronal permet aux interlocuteurs de s'accommoder par un processus dit d'« entraînement » au rythme de la parole de l'autre (Peelle & Davis, 2012). La fréquence oscillatoire du débit de parole de l'autre étant alors simulée de manière endogène, l'auditeur peut ainsi facilement émettre des prédictions sur le moment où le tour de parole de son interlocuteur va prendre fin et sur le moment où son tour, en tant que locuteur, va devoir commencer ; la programmation temporelle de sa parole étant conditionnée par ces processus anticipatoires (Garrod & Pickering, 2015).

De récentes études utilisant la technique des mouvements oculaires (Casillas & Frank, 2013) ont montré que les enfants sont capables, dès trois ans, d'anticiper l'occurrence temporelle des tours de parole de deux interlocuteurs en train de dialoguer ; ce phénomène étant plus particulièrement marqué lors de la présence d'énoncés hautement prévisibles tels que les paires de questions-réponses. En production, les études s'intéressant au développement de cette compétence s'attachent majoritairement à évaluer la qualité de synchronisation de dyades mère-bébé à un niveau préverbal (Jaffe et al., 2001). Il semble cependant intéressant de coupler les recherches récentes en perception chez l'enfant d'âge verbal à des tâches en production, permettant de mesurer leurs capacités de convergence temporelle dans le cadre d'une interaction. Pour ce faire, nous avons créé une tâche dans laquelle un enfant dénomme des images en alternance avec un partenaire virtuel. Ce type de paradigme nous a permis de manipuler deux paramètres temporels : le tempo de l'alternance et la constance de la rythmicité des mots échangés. Ces deux facteurs nous permettent de mesurer la convergence de l'enfant vers une temporalité interactionnelle imposée.

## 2 Matériel et Méthode

### 2.1 Population

Nous avons recruté un groupe de 16 enfants, âgés de 5 et 6 ans (étendue = 52 mois à 71 mois ; moyenne d'âge = 65 mois ; écart-type= 5 mois), de langue maternelle française. Ce groupe était composé de 10 filles et 6 garçons. Les enfants porteurs de troubles visuels, de troubles du langage ou de troubles auditifs ont été exclus de l'étude. Ces enfants francophones étaient tous scolarisés dans une école du centre de Marseille en classe de moyenne et de grande section de maternelle. Nous avons reçu le consentement éclairé de tous les parents avant de commencer l'étude. Les enfants ont tous reçu un livre en remerciement de leur participation.

### 2.2 Stimuli

Pour cette tâche de dénomination, nous avons réalisé une sélection de 160 images comportant 80 mots monosyllabiques et 80 mots bisyllabiques issues de la banque de d'images BD21 (Cannard et al., 2006) en fonction de leur taux de dénomination pour un âge d'acquisition allant de 3 à 8 ans. La moyenne du taux de réussite de dénomination pour les mots monosyllabiques est de 93,9, écart-type= 5,3 et la moyenne du taux de réussite de dénomination pour les mots bisyllabiques est de 98,8, écart-type= 1,6. Le lexique représenté par ces images était composé de 17 catégories taxonomiques : aliment, animal, bijou, fourniture scolaire, habitation, instrument de musique, jouet, meuble, outil, partie du corps, paysage, personne, plante, ustensile de cuisine, véhicules, vêtements, autre. A partir des images sélectionnées, nous avons créé des paires d'images : la première image de la paire étant destinée au partenaire virtuel (PV) et la deuxième à l'enfant (figure 1). La composition de ces paires était différente en fonction des listes de mots créées. Dans les listes à rythmicité constante (condition *match*), les paires étaient composées de mots avec un nombre de syllabes identique (i.e. dans certaines listes toutes les paires étaient composées de mots monosyllabiques et dans d'autres, toutes les paires étaient composées de mots bisyllabiques). Dans les listes à rythmicité aléatoire (condition *mismatch*), les paires étaient composées de mots avec un nombre de syllabes aléatoirement identique ou différent (i.e. certaines paires étaient composées de mots monosyllabiques ou bisyllabiques et, à l'intérieur de la même liste, certaines paires étaient composées de mots avec un nombre de syllabes différent : par exemple, le PV dénommait un mot monosyllabique alors que l'enfant avait à dénommer un mot bisyllabique). Nous avons ainsi créé 8 listes de 10 paires d'images chacune selon cette procédure : 4 listes *match* et 4 listes *mismatch*. Pour éviter des effets d'amorçage sémantique, catégoriel ou phonétique, nous avons pris soin de ne jamais mettre deux mots de la même catégorie, commençant par le même phonème ou appartenant à la même catégorie dans une paire d'images ou dans deux paires successives.

Les images destinées au PV ont été dénommées à débit normal et enregistrées à partir d'une voix féminine, dans une chambre sourde à un échantillonnage de 44.1 kHz.

Afin de générer la perception d'une rythmicité dans l'échange, nous avons extrait, grâce à un algorithme développé par Cummins & Port (1998), tous les points acoustiques déterminés par notre système perceptif comme étant le début de chaque syllabe (« p-centers ») des mots enregistrés et nous les avons fait délivrés, via le logiciel Presentation (Neurobehavioral System) à intervalles réguliers : soit toutes les 3200 ms (condition *lente* du tempo de l'alternance), soit toutes les 2600 ms (condition *rapide*).

L'ordre de présentation des listes à l'intérieur des 2 modalités du facteur Vitesse d'alternance, ainsi que l'ordre de présentation des modalités du facteur Rythmicité lors du test, ont été contrebalancés. Toutes les listes de paires d'images étaient précédées d'un exemple de 3 échanges entre un enfant et le PV en « parfaite » alternance afin de simuler la situation expérimentale et de familiariser l'enfant avec la rythmicité et le tempo de la liste à venir. Ainsi, pour ces 3 paires de mots, nous avons placé le p-center de l'enfant virtuel à la moitié de l'intervalle séparant les 2 p-centers successifs du partenaire virtuel (figure 2) c'est-à-dire à 1600 ms pour la modalité *lente* et 1300 ms pour la modalité *rapide*.

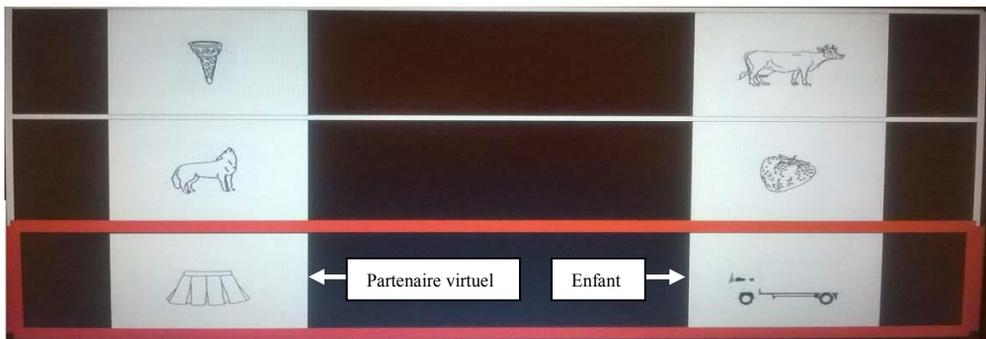


FIGURE 1- Capture d'écran de la tâche de dénomination. Les images de gauche sont dénommées par le partenaire virtuel et celles de droite par l'enfant. Le cadre rouge défile régulièrement sur l'écran au fur et à mesure de la succession des paires d'images.

Ces stimuli sonores ont été délivrés en utilisant une carte son de marque Creative via une enceinte de marque Sony. Les dénominations des images des enfants ont été enregistrées avec un micro-casque de marque Sennheiser. Les paires d'images ont été délivrées via un ordinateur portable de marque DELL, résolution 1366 X 768 (figure 1).

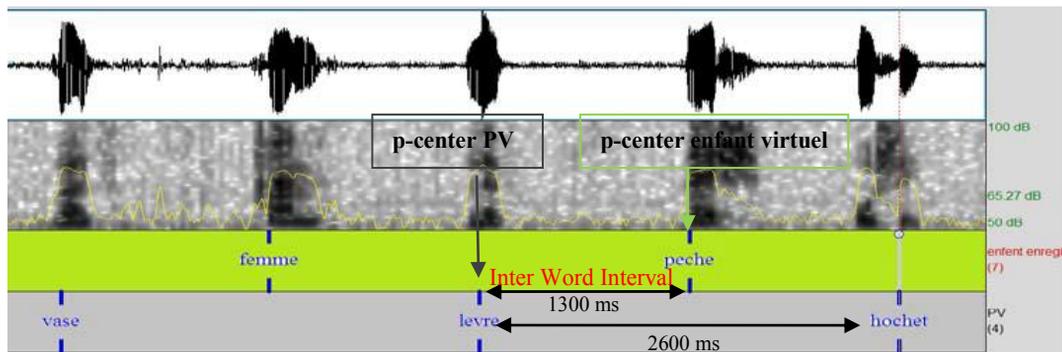


FIGURE 2- Textgrid (Praat) montrant le placement régulier des p-centers sur des paires de mots préenregistrés (PV+enfant virtuel) selon une rythmicité d'alternance « parfaite » délivrée avant chaque liste.

## 2.3 Procédure

Avant le test, une étape de familiarisation avec les images de chaque liste ainsi qu'avec le principe de l'alternance a été effectuée afin de s'assurer que l'enfant serait en mesure de réaliser la tâche convenablement, qu'il reconnaîtrait l'image de l'objet, que le lexique serait connu et que l'intelligibilité serait suffisante pour assurer les mesures et les analyses.

Lors du test, les images ont été présentées à distance et à intensité confortables en fonction de chaque enfant. La consigne était de regarder l'écran d'ordinateur et de dénommer la deuxième image de la paire à la suite de la dénomination de la première image par le PV. Il était explicitement demandé que la dénomination soit réalisée à la même vitesse que l'enfant qu'il venait d'entendre parler. L'enfant était encouragé entre chaque liste.

La succession des images n'a pas été interrompue même si l'enfant ne se souvenait plus du lexique représenté par l'image afin de ne pas rompre la rythmicité créée par l'alternance. Chaque enfant a été testé individuellement par le même expérimentateur dans une salle d'étude d'une école du centre de Marseille.

## 2.4 Mesures

Afin d'analyser si l'enfant s'améliorait dans la perception et la réalisation de la régularité alternée en fonction des conditions, nous avons mesuré l'espace séparant le p-center de chaque mot du PV et celui de l'enfant pour chaque paire d'images. Cet intervalle, appelé Inter-Word-Interval (IWI) (figure 2 en rouge) servait de base pour le calcul de l'accommodation. Les p-centers ont été extraits de manière semi-automatique (Cummins & Port, 1998), leur localisation inspectée et, si nécessaire, corrigée manuellement dans le logiciel PRAAT (Boersma, 2002). Les mots précédés par un « euh » d'hésitation, un article, un bredouillement (duplication du 1er phonème) ont été exclus des analyses. De même, nous avons éliminé les essais (151 sur 1280) lorsqu'un temps de latence trop long de la part de l'enfant a généré un chevauchement de parole avec le mot suivant du partenaire virtuel.

## 2.5 Analyses statistiques

A partir de notre mesure, l'*Inter-word-Interval (IWI)*, qui est l'intervalle de temps qui sépare le p-center du PV de celui de l'enfant, nous avons conduit deux types d'analyses. Nous avons tout d'abord réalisé, à l'aide du logiciel Statistica, une analyse de la variance à mesures répétées avec deux facteurs intra-sujets : Vitesse (*lent* versus *rapide*) et Rythmicité (*match* versus *mismatch*), le seuil de significativité étant fixé à  $\alpha=0.05$ . Nous avons ensuite conduit, à l'aide de la boîte à outils CircStat du logiciel Matlab, des statistiques circulaires descriptives. Ce type d'analyses, nous a permis de transposer les réponses de chaque enfant dans un espace vectoriel dont la longueur du cercle représente un cycle oscillatoire de 2600 ms en condition *rapide* et 3200 ms en condition *lente*. Le point 0 (ou 360°) représente le placement du p-center du PV et le point à l'antiphase du point 0, c'est à dire à 180°, représente le moment « attendu » du placement du p-center de l'enfant. Les points situés dans l'espace supérieur à ce point traduisent les réponses trop précoces et les points situés dans l'espace inférieur les réponses trop tardives par rapport à une parfaite régularité rythmique entre les productions des locuteurs. La moyenne de tous ces points a ensuite été représentée par un vecteur d'angle  $\Theta$  (exprimé en radians) et de longueur  $r$ . L'angle de ce vecteur permet de déterminer la précision des réponses de l'enfant par rapport au moment attendu ou encore le niveau d'*asynchronie* : plus l'angle est grand, plus les réponses de l'enfant sont éloignées du moment attendu. La longueur du vecteur permet de déterminer la consistance de

ses réponses, c'est à dire leur degré de variabilité : plus le vecteur est long, plus la *consistance des réponses* de l'enfant est importante. Nous avons tout d'abord réalisé le test d'hypothèses de Rayleigh afin de contrôler le caractère non aléatoire de la distribution des réponses autour du cercle pour chaque enfant, le seuil de significativité étant fixé à  $\alpha=.05$ . La distribution relevée dans chaque condition ne révélant pas de distribution au niveau du hasard ( $p<.05$ ), nous avons pu procéder à l'analyse de l'angle et de la longueur du vecteur. A partir de la moyenne des angles  $\Theta$ , obtenue en radians dans chaque condition et pour chaque enfant, nous avons transformé ces radians en millisecondes puis réalisé une analyse de la variance à mesures répétées avec les facteurs intra-sujets *Vitesse* et *Rythmicité*. A partir de la moyenne de la longueur des vecteurs obtenue dans chaque condition, nous avons réalisé une analyse de la variance avec les mêmes facteurs intra-sujets que pour les angles (*Vitesse* et *Rythmicité*).

### 3 Résultats

#### 3.1 Prise en compte de la vitesse et influence sur la précision des réponses

La figure 3 montre un effet significatif de vitesse. Les enfants ont en effet placé leurs mots en

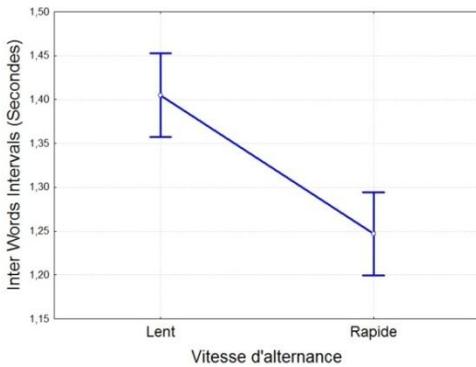


FIGURE 3 - Intervalles en secondes entre les p-center de l'enfant et ceux du partenaire virtuel en fonction de la vitesse d'alternance. Les barres verticales représentent les intervalles de confiance à 95%.

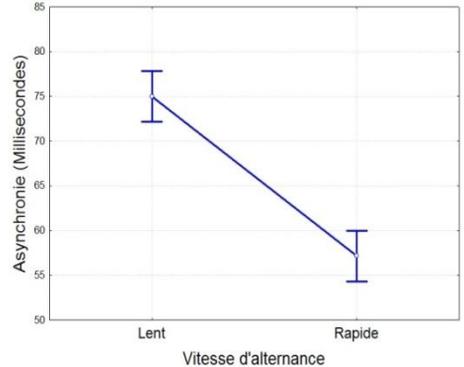


FIGURE 4 - Degré d'asynchronie en millisecondes (mesure d'imprécision) du placement du p-center de l'enfant par rapport au moment attendu en fonction des conditions de vitesse de l'alternance. Les barres verticales représentent les intervalles de confiance à 95%.

fonction du tempo d'alternance imposé, conservant ainsi une alternance plus rapide dans la condition *rapide* par rapport à la condition *lent*  $F(1, 15) = 24.988, p<.000$ . Les mesures effectuées à partir des statistiques circulaires sur la *grandeur de l'angle* révèlent également un effet principal de vitesse sur l'*asynchronie*. Les enfants sont moins précis dans la condition *lent* que dans la condition *rapide* :  $F(1, 15)=89.710, p<.000$  (figure 4).

#### 3.2. Consistance des réponses : effet de vitesse et influence de la régularité de rythmicité

Les mesures effectuées à partir de la *longueur du vecteur* révèlent que les enfants répondent de manière plus *consistante* dans la condition où la vitesse d'alternance est *rapide* comparé à la condition *lente*  $F(1, 15)=6.852, p=.019$  et que leurs réponses sont plus consistantes pour les listes à *rythmicité constante* plutôt qu'*aléatoire*  $F(1,15)=4.093, p=.061$  (effet marginal) (figure 5).

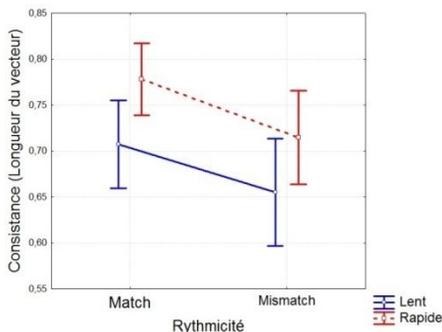


FIGURE 5 - Degré de consistance du placement des p-centers en fonction de la vitesse d'alternance et de la rythmicité des listes de mots. Les barres verticales représentent les intervalles de confiance à 95%.

## 4 Discussion et Conclusion

### 4.1 Effet de la vitesse de dénomination sur le système perceptivo-moteur

Les résultats obtenus à partir de ce nouveau paradigme montrent non seulement que les enfants de 5-6 ans sont capables d'adapter leur vitesse de dénomination à celle imposée par un partenaire virtuel mais également que notre tâche est pertinente pour tester l'accommodation temporelle chez des enfants de cette tranche d'âge. En outre, le fait que les enfants répondent au-delà du hasard dans toutes les conditions et qu'ils soient systématiquement en avance par rapport au moment de dénomination attendu nous prouve que les données récoltées ne sont pas du fait de la latence de dénomination mais bien de leur convergence temporelle. De plus, il est intéressant de noter que les enfants sont plus proches du temps attendu en condition *rapide* qu'en condition *lente*. Cette précision temporelle de la dénomination de l'enfant par rapport au PV est le facteur qui conditionne la réalisation d'une régularité parfaite de l'alternance. Cette précision temporelle accrue en condition de vitesse rapide génère une alternance des mots (PV+enfant) toutes les 1300 ms contre 1600 ms en vitesse lente. Dans les quelques données connues sur le développement du timing du tour de parole chez l'enfant, les études ont pu montrer qu'à 3 ans, les enfants qui échangent entre eux présentent en moyenne des transitions de tour de parole situées autour de 1500 ms mais qu'à 4 ans, ils se rapprochent du timing des adultes situé entre 20 et 1000 ms (Casillas, 2014). On peut alors penser que la vitesse proposée en condition *rapide* est plus proche d'une vitesse « naturelle » d'échange d'un enfant de 5 ans que la vitesse lente. D'autre part, l'augmentation de la consistance des réponses à cette vitesse de dénomination, suggère que dans la condition *rapide*, le système sensori-moteur de l'enfant arrive mieux à se synchroniser sur le stimulus de parole. Les études portant sur la synchronisation sensorimotrice telles que celles utilisant la technique de « tapping », montrent que la synchronisation avec un métronome devient

difficile dès lors qu'on se rapproche de 1800 ms (MacDorman, 1962, cité dans Repp & Doggett, 2007). La condition lente générant une succession des p-centers toutes les 1600 ms, elle place le système perceptivo-moteur des enfants dans une condition plus proche de cette limite de perception de l'isochronie que la condition *rapide* (1300 ms).

## 4.2 Effet de la rythmicité sur le système perceptivo-moteur

En faisant varier le nombre de syllabes des mots échangés entre le partenaire virtuel et l'enfant, nous souhaitions analyser si la perturbation de la rythmicité de l'alternance pouvait avoir une influence sur la régularité de dénomination des enfants ; ceci nous renseignant non seulement sur la capacité des enfants de 5-6 ans à percevoir une rythmicité induite dans une alternance de parole mais également à la reproduire. Les données issues des statistiques circulaires mettent en exergue la sensibilité des enfants à cette variation confirmant ainsi notre hypothèse. La plus grande consistance des réponses retrouvées dans la condition où la rythmicité des paires de mots est demeurée constante (condition *match*) montre que les enfants sont sensibles à une rythmicité induite dans un rythme de parole mais aussi que celle-ci a une conséquence sur la programmation de leur parole : les enfants sont plus réguliers lorsque la rythmicité des paroles échangées entre les locuteurs est régulière comparée à irrégulière. Ainsi, si l'on appréhende ce résultat selon les mécanismes de la dynamique oscillatoire neuronale, on peut penser qu'une régularité dans la rythmicité des paires de mots échangées permet plus facilement l'émergence du phénomène d'« *entrainement* » c'est-à-dire du couplage de phase entre le stimulus de parole et les populations neuronales du système auditif (Large & Jones, 1999) ; l'émergence plus évidente d'une métrique dans la condition de rythmicité constante étant probablement à la base de ce phénomène facilitateur (Nozaradan et al., 2012). En outre, il est également possible que cette condition ait favorisé un couplage de phase entre les systèmes auditifs et moteurs (Large et al., 2015), augmentant la génération de prédictions temporelles sur les p-centers à venir ainsi que la régularité de la distribution des dénominations.

## 4.3 Dénomination alternée - un paradigme utile pour la recherche sur les capacités d'interaction chez l'enfant?

Si l'on transpose les effets obtenus avec ce nouveau paradigme dans le cadre d'une interaction naturelle, cela signifierait que l'enfant parvient, dans certaines conditions de tempo conversationnel, à prendre son tour de parole au moment *le plus opportun* dans la conversation c'est-à-dire en ne chevauchant pas la parole de son interlocuteur, et en ne laissant pas trop de délai entre le tour de parole de celui-ci et le sien. En outre, les effets observés suite à la manipulation de la rythmicité de l'alternance montrent que la régularité de la parole de l'autre influence la capacité de l'enfant à s'accommoder et à demeurer régulier dans la programmation de sa parole.

Certaines conditions temporelles de l'alternance ont donc favorisé chez l'enfant le phénomène d'anticipation utilisé dans la gestion des tours de parole (Pickering & Garrod, 2015). Ceci nous permet de penser que grâce à ce type de paradigme et à la manipulation de facteurs temporels tels que la rythmicité de parole, des compétences temporelles sous-jacentes à la fluidité de la dynamique interactionnelle pourront désormais être évaluées chez des populations cliniques d'enfants souffrant de troubles du langage. En outre, l'effet de la manipulation de la rythmicité sur la régularité de la production de parole nous conforte dans le fait que des techniques thérapeutiques, telles que la musique, visant au développement du traitement temporel auditif, devraient plus systématiquement être proposées dans la prise en charge des troubles du langage chez l'enfant.

## Remerciements

Cette étude a été réalisée grâce au soutien financier du Brain and Language Research Institute (BLRI, ANR-11-LABX-0036), Laboratoire Parole et Langage - CNRS & Université d'Aix-Marseille.

## Références

- BOERSMA, P. (2002). Praat, a system for doing phonetics by computer. *Glott international* 5, 341-345.
- BRANIGAN, H. P., PICKERING, M. J., CLELAND, A. A. (2000). Syntactic co-ordination in dialogue. *Cognition* 75, B13–B25.
- CANNARD, C., BLAYE, A., SCHEUNER, N., & BONTHOUX, F. (2005). Picture naming in 3-to 8-year-old French children : Methodological considerations for name agreement. *Behavior Research Methods* 37, 417-425.
- CASILLAS, M. (2014). Pragmatic development in first language acquisition. Amsterdam/Philadelphia : Matthews, D.
- CASILLAS, M., FRANK, M. C. (2013). The development of predictive processes in children's discourse understanding. In *CogSci 2013 The 35th annual meeting of the Cognitive Science Society*, 299-304.
- CUMMINS, F., PORT, R. (1998). Rhythmic constraints on stress timing in English. *Journal of Phonetics* 26, 145-171.
- GARROD, S., PICKERING, M. J. (2015). The use of content and timing to predict turn transitions. *Frontiers in Psychology* 6, 1–12.
- JAFFE, J., BEEBE, B., FELDSTEIN, S., CROWN, C. L., JASNOW, M. D., ROCHAT, P., STERN, D. N. (2001). *Rhythms of dialogue in infancy : Coordinated timing in development*. Temple University : W. F., Overton.
- KRIVOKAPIC, J. (2013). Rhythm and convergence between speakers of American and Indian English. *Laboratory Phonology* 4, 39–65.
- LARGE, E. W., HERRERA, J. A., VELASCO, M. J. (2015). Neural Networks for Beat Perception in Musical Rhythm. *Frontiers in Systems Neuroscience* 9.
- LARGE, E. W., JONES, M. R. (1999). The dynamics of attending : How people track time-varying events. *Psychological Review* 106, 119–159.
- NOZARADAN, S., PERETZ, I., MOURAUX, A. (2012). Selective neuronal entrainment to the beat and meter embedded in a musical rhythm. *The Journal of Neuroscience* 32, 17572–81.
- PARDO, J. S. (2006). On phonetic convergence during conversational interaction. *The Journal of the Acoustical Society of America*, 119, 2382-2393.
- PEELLE, J. E., DAVIS, M. H. (2012). Neural Oscillations Carry Speech Rhythm through to Comprehension. *Frontiers in Psychology* 3, 1-17.
- REPP, B. H., DOGGETT, R. (2007). Tapping to a very slow beat : A comparison of musicians and nonmusicians. *Music Perception* 24, 367–376.
- SCOTT, S. K., MCGETTIGAN, C., EISNER, F. (2009). A little more conversation, a little less action-candidate roles for the motor cortex in speech perception. *Nature Reviews Neuroscience* 10, 295-302.
- WILSON, M., WILSON, T. P. (2005). An oscillator model of the timing of turn-taking. *Psychonomic Bulletin & Review* 12, 957–68.