

# Détection et quantification automatisée de poussées de croissance

Frédéric Dandurand, Ph.D.

Colloque en Méthodes Quantitatives  
et Sciences Humaines

Université de Montréal, 7 juin 2010

# Plan

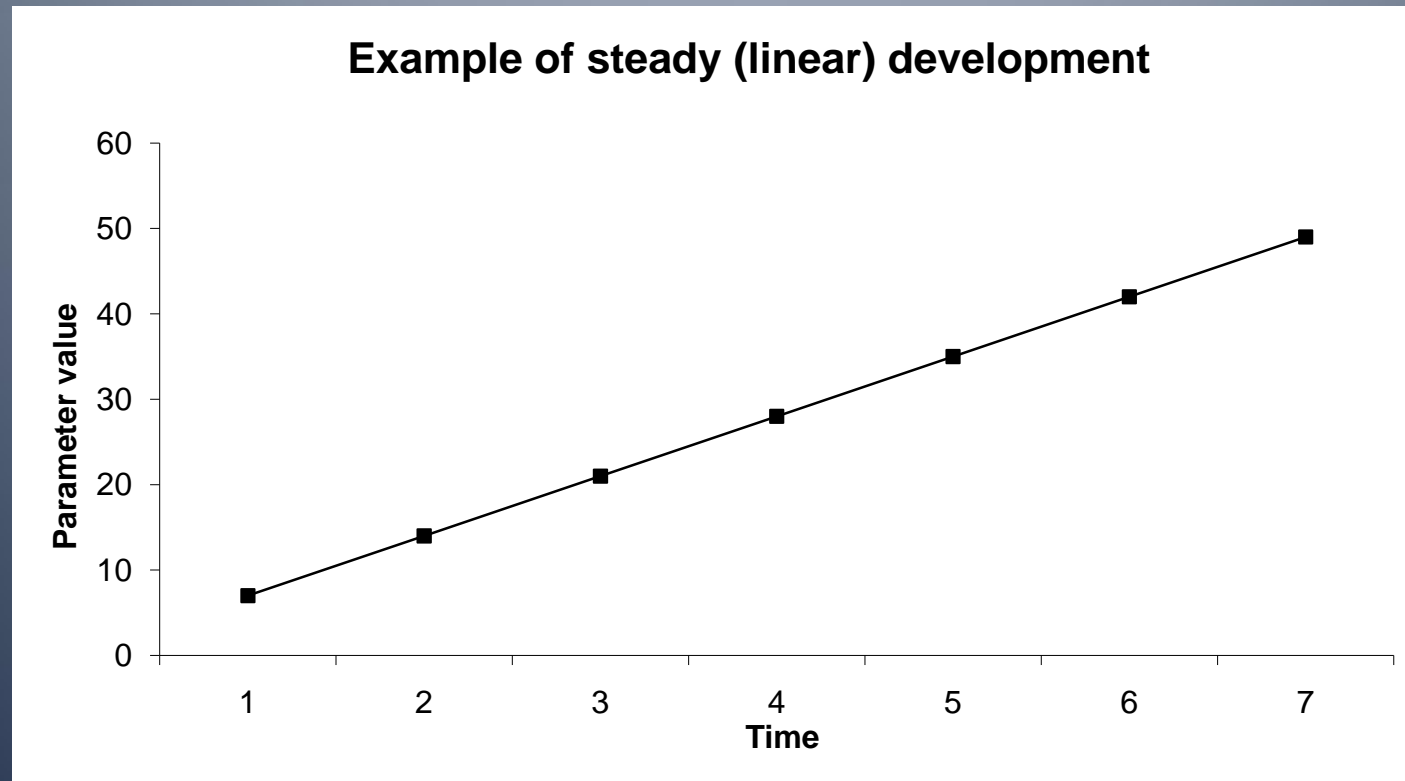
- Défis de l'analyse de données longitudinales
- Nouvel outil pour détecter et quantifier les poussées de croissance
- Exemples
  - Croissance physique (taille)
  - Apprentissage des mots (taille du lexique)
- Perspectives
  - Application aux données temporelles (imagerie)
  - Stades de développement en psychologie

# Croissance

- Phénomène important et commun
  - Systèmes biologiques
  - Croissance psychologique
  - Economie (PIB), etc.
- Question fondamentale -- caractérisation de la croissance: linéaire ou pas?

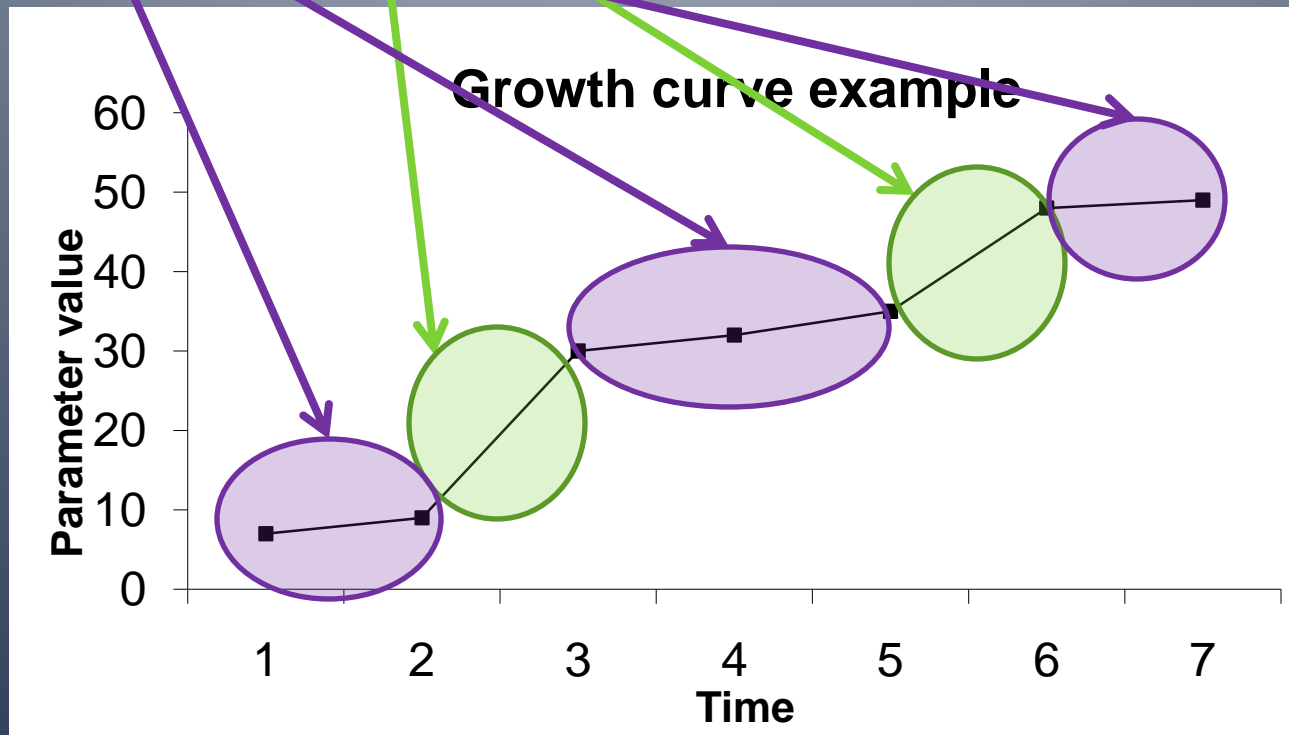
# Croissance linéaire

- Taux de croissance (pente, dérivée première) constant



# Croissance non-linéaire

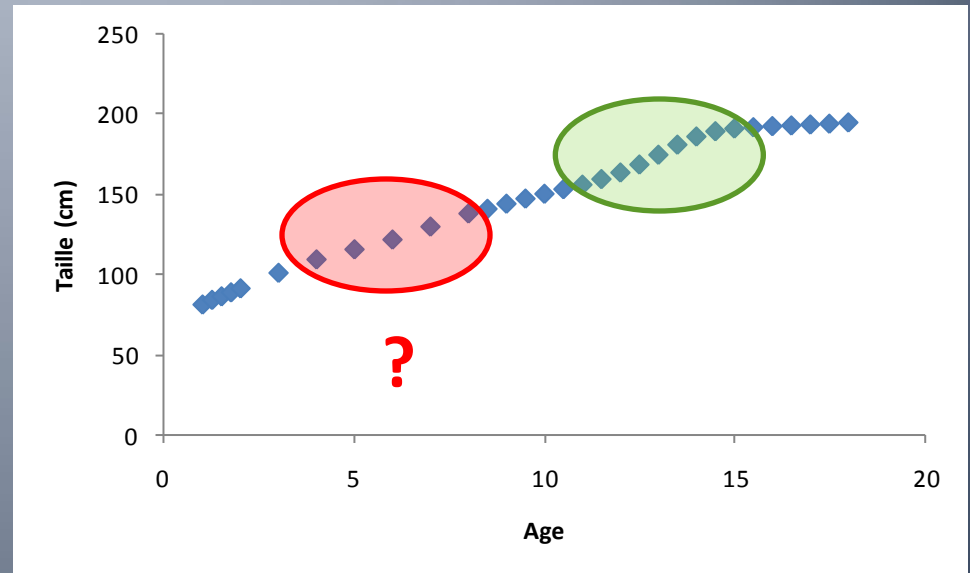
- Taux de croissance non-constant
- Plateaux et poussées



# Technique exploratoire : Inspection visuelle

- Où sont les poussées?
- Combien de temps durent-elles?
- Difficile de marquer le début et la fin
- Bonus: Données expérimentales bruitées
  - Bruit peut causer des augmentations du taux de croissance
  - Reconnaître les poussées réelles, fiables

Taille d'un garçon (Berkeley growth study, 1954)



- Sommaire:
  - Arbitraire et ouverte à l'interprétation
  - Manque de rigueur

# Solution: Automatic Maxima Detection

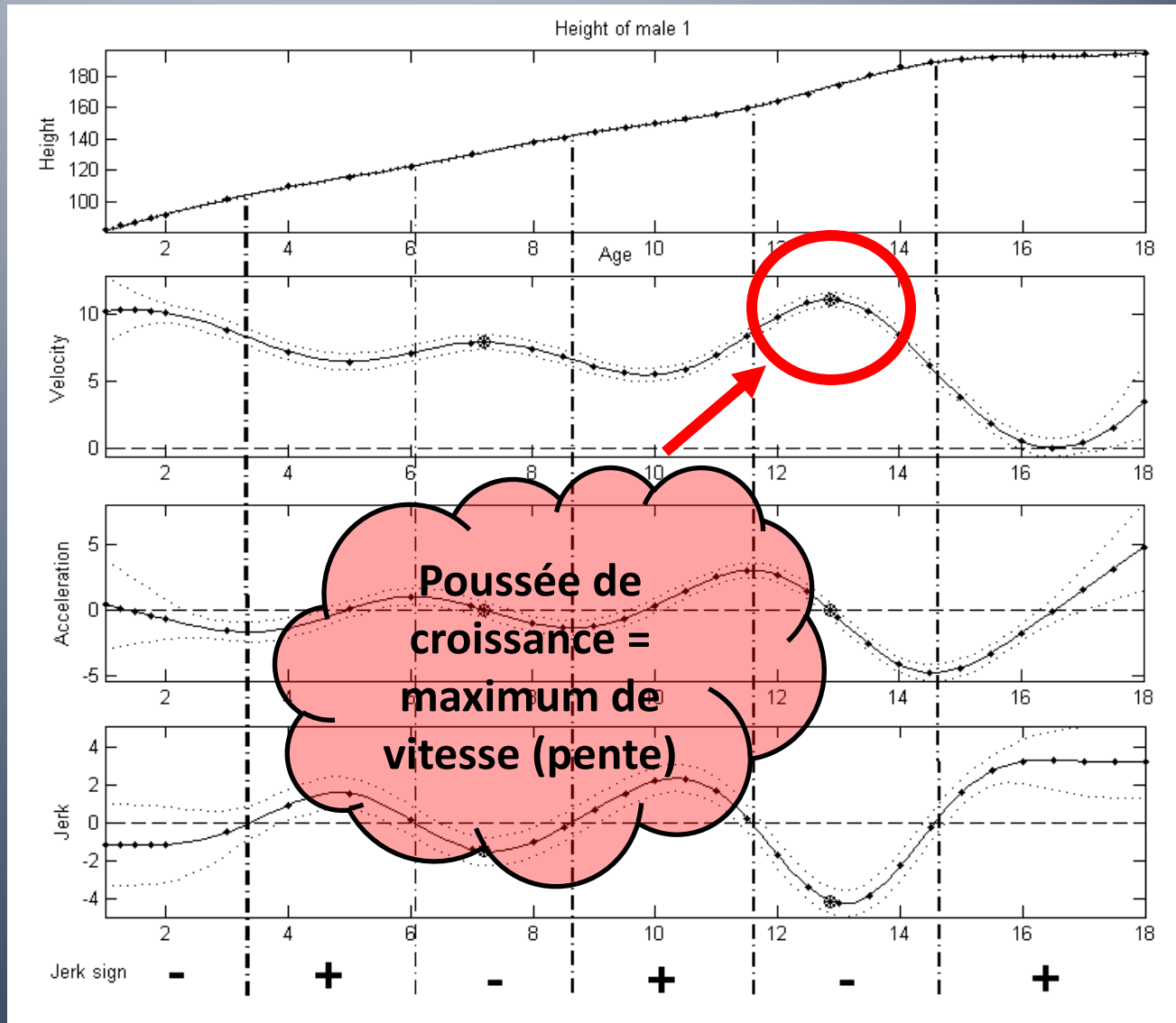
- Approximer les données observées (mesurées) par une fonction continue
  - Fonction lisse de type “spline”
  - Approximation des dérivées
  - Calcul des intervalles (bandes) de confiance (ex.: 95%)
- Automatic Maxima Detection (AMD)
  - Identifier les poussées significatives
  - Quantifier ces poussées
  - Outils de visualisation

Note: Implémentation (non-restrictive) utilisant  
Functional Data Analysis (Ramsay & Silverman, 2005)

# Exemple – Taille physique

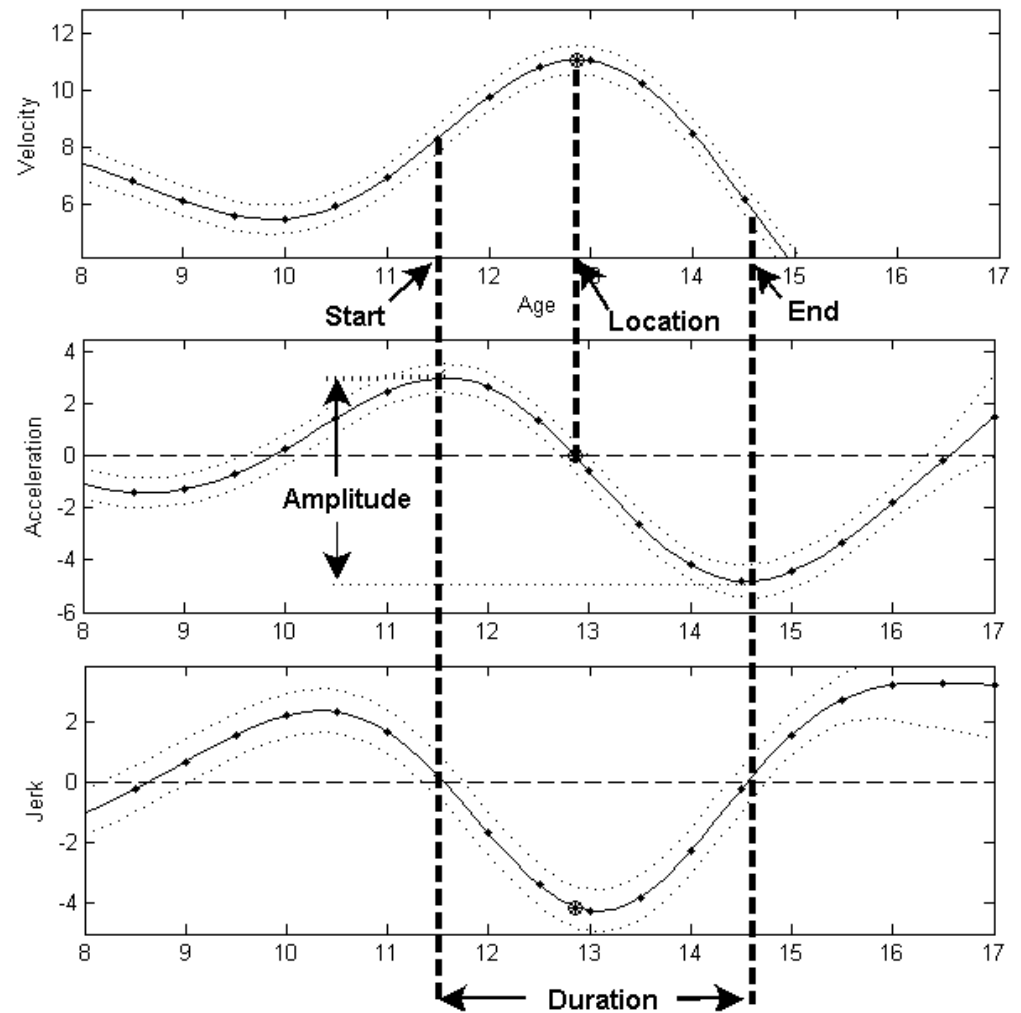
Fonction et ses  
3 premières  
dérivées

Passages par  
zéro, maxima,  
minima



# Quatre mesures de poussées

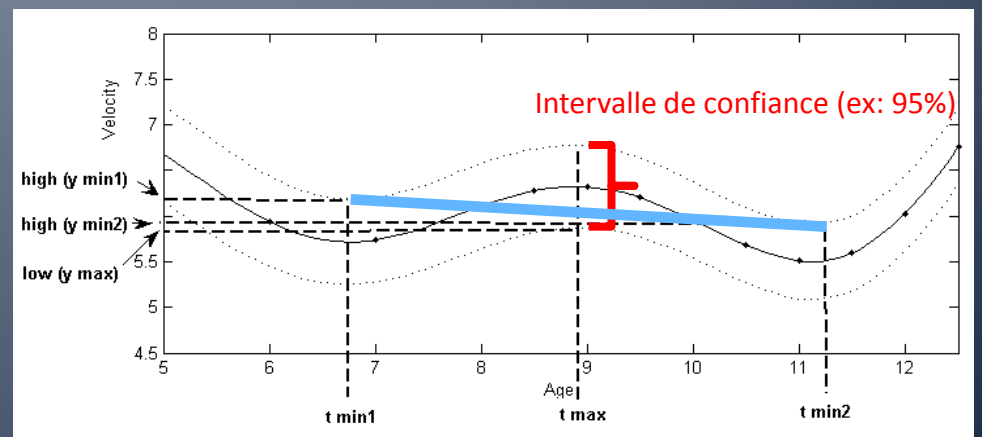
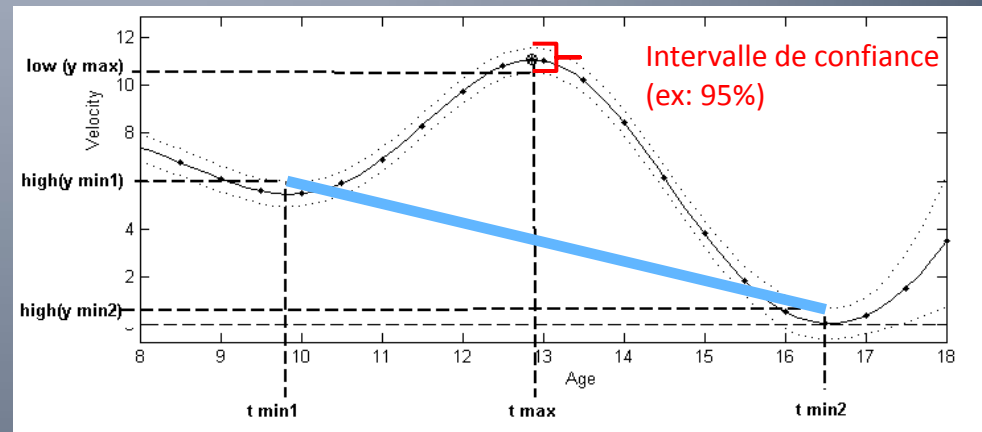
- Position (location)
- Début (start)
- Durée (duration)
- Amplitude



# Poussée statistiquement significative?

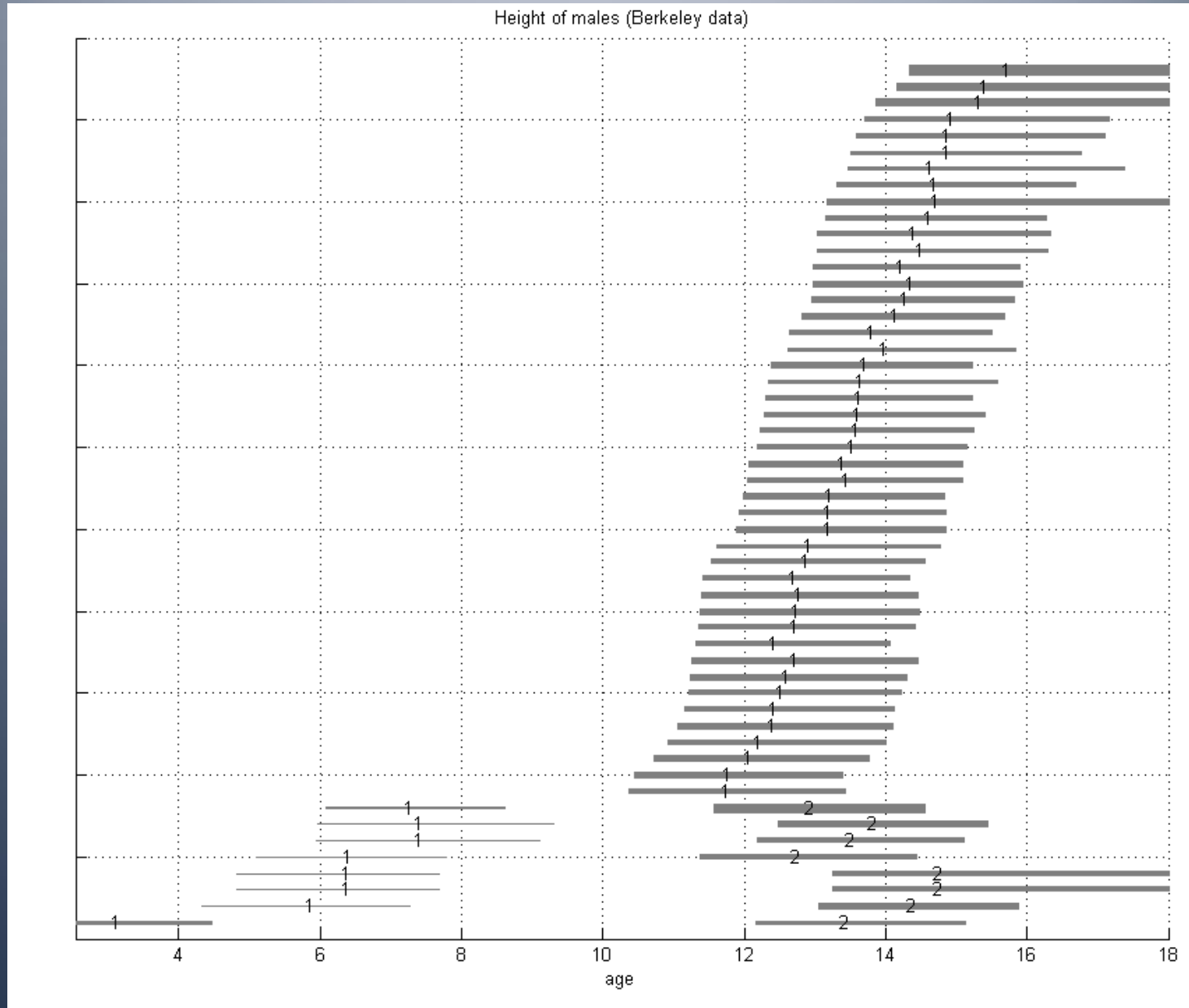
$H_0$ : Une droite (donc sans maximum) contenue dans l'intervalle de confiance rend compte des données

Rejeter  $H_0$ :  
Poussée significative



On ne peut pas rejeter  $H_0$ :  
Non-significative

# Résultats : Berkeley growth - garçons



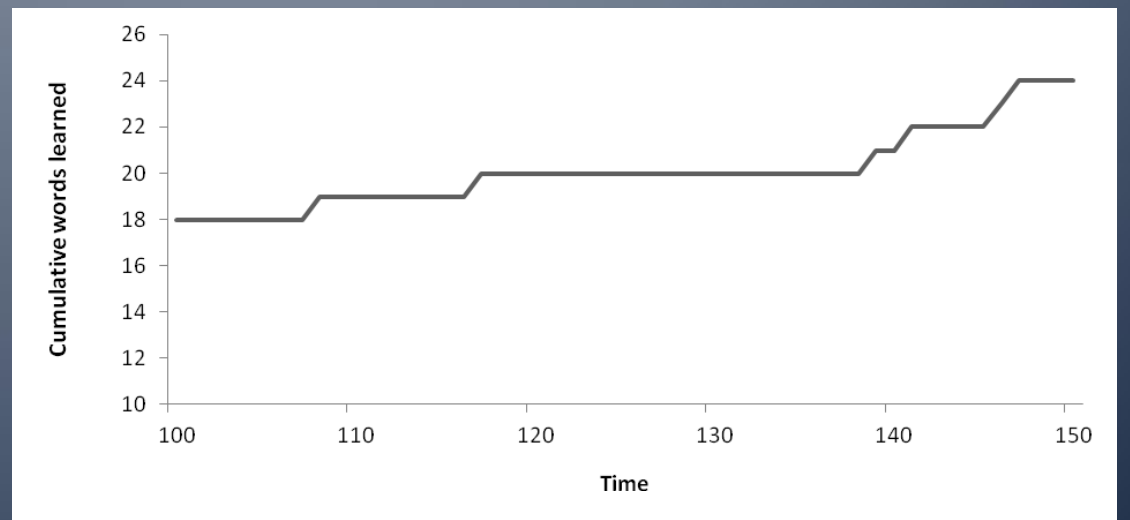
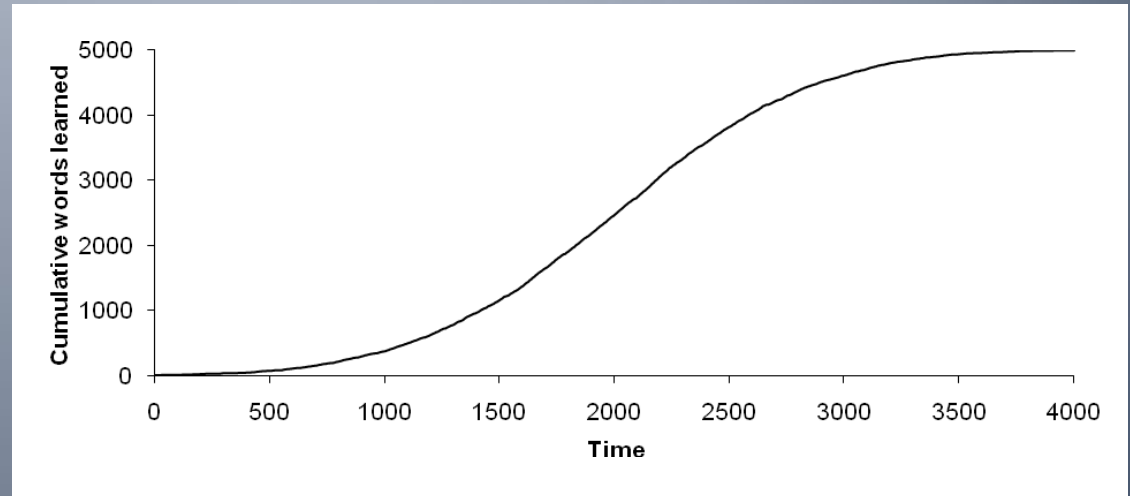
# Poussée de croissance à l'adolescence

- Soumettre les mesures générées par AMD à des techniques classiques d'analyse statistique
- Exemple: ANOVA pour comparer les poussées de croissances chez les filles et les garçons
  - Réplique des résultats connus et établis
    - Début & position (âge): filles < garçons
    - Durée (années): filles < garçons
    - Amplitude: filles < garçons

# Modèle de croissance du vocabulaire

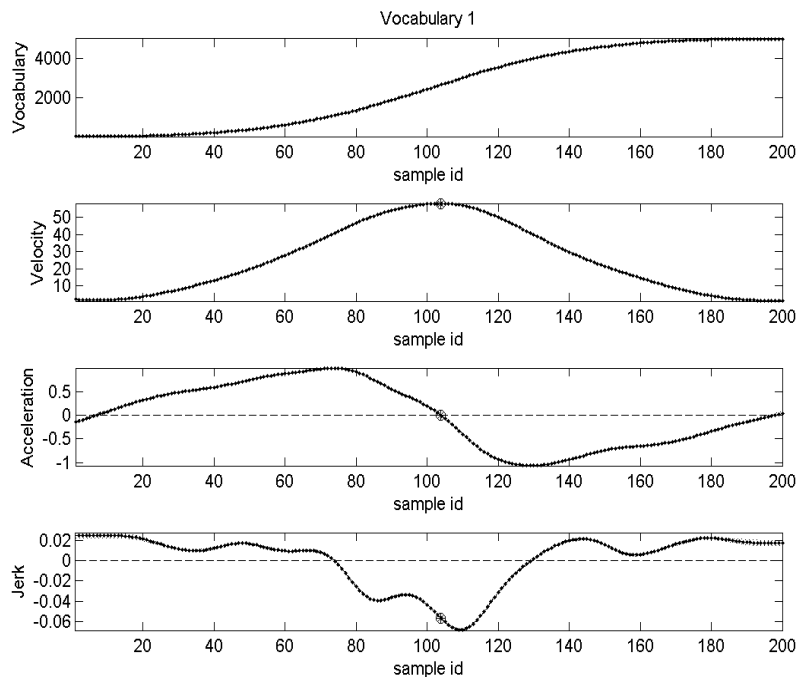
- Nombre de mots appris en fonction du temps
- Seuls des mots complets peuvent être appris

(Mitchell, & McMurray, 2009)

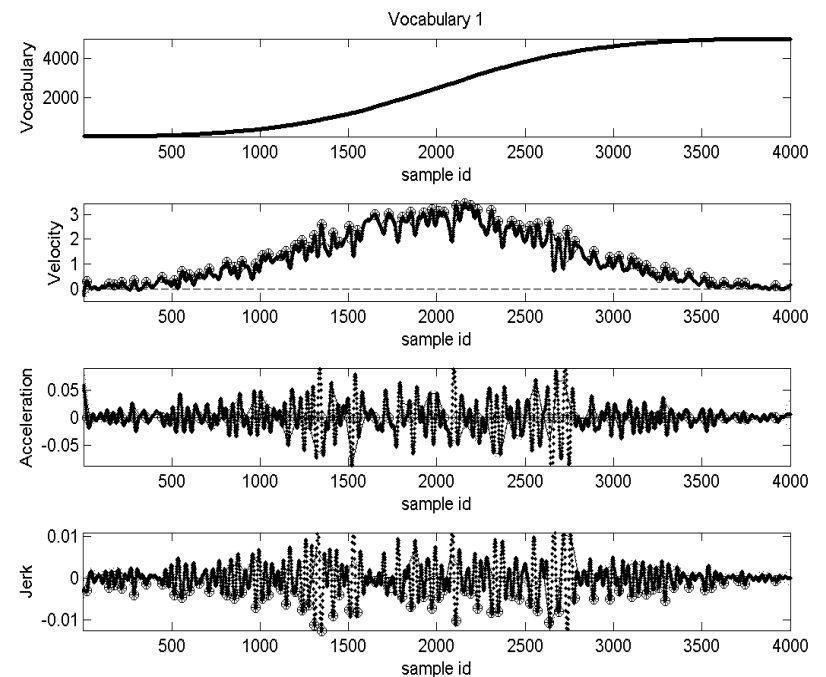


# Poussées globales et locales

Lissage fort

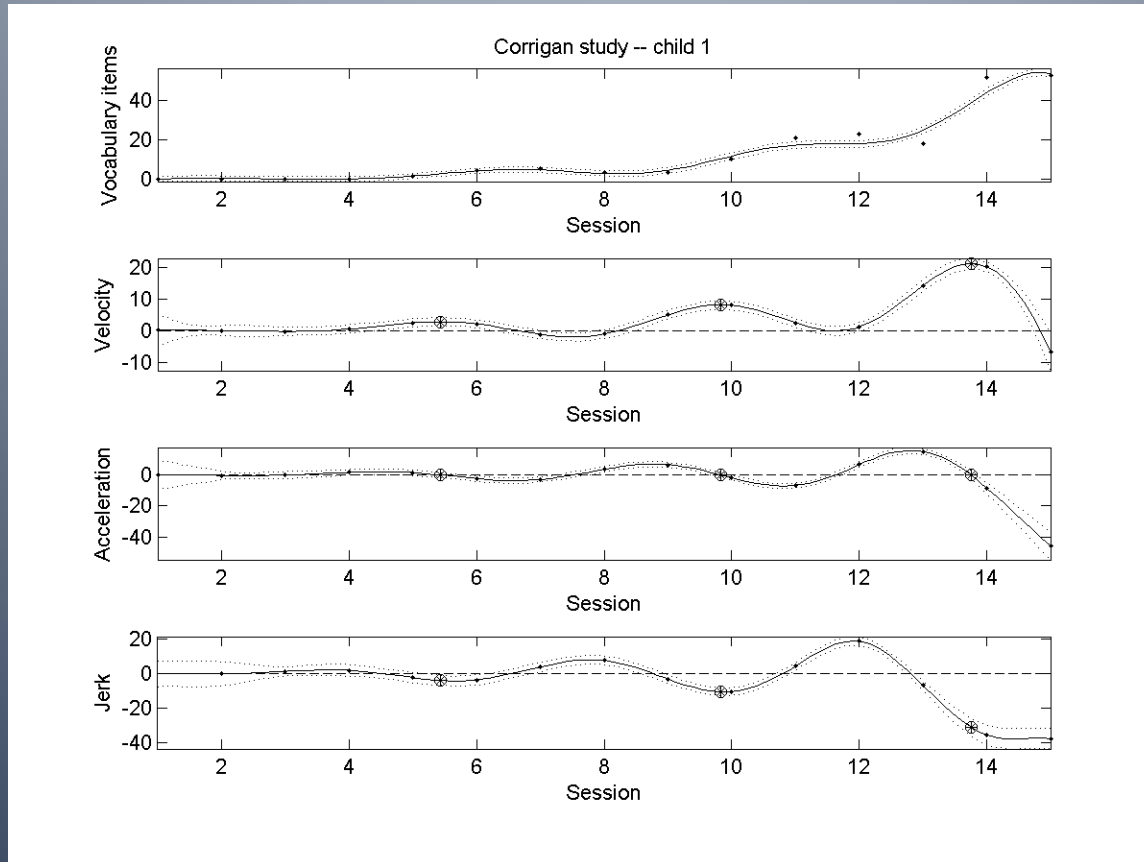


Lissage faible



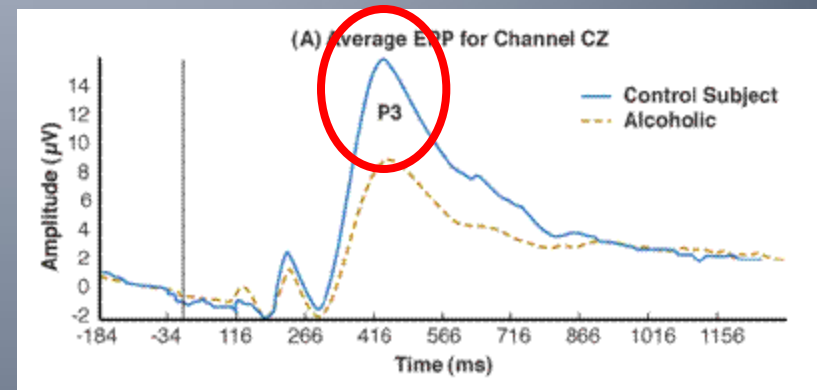
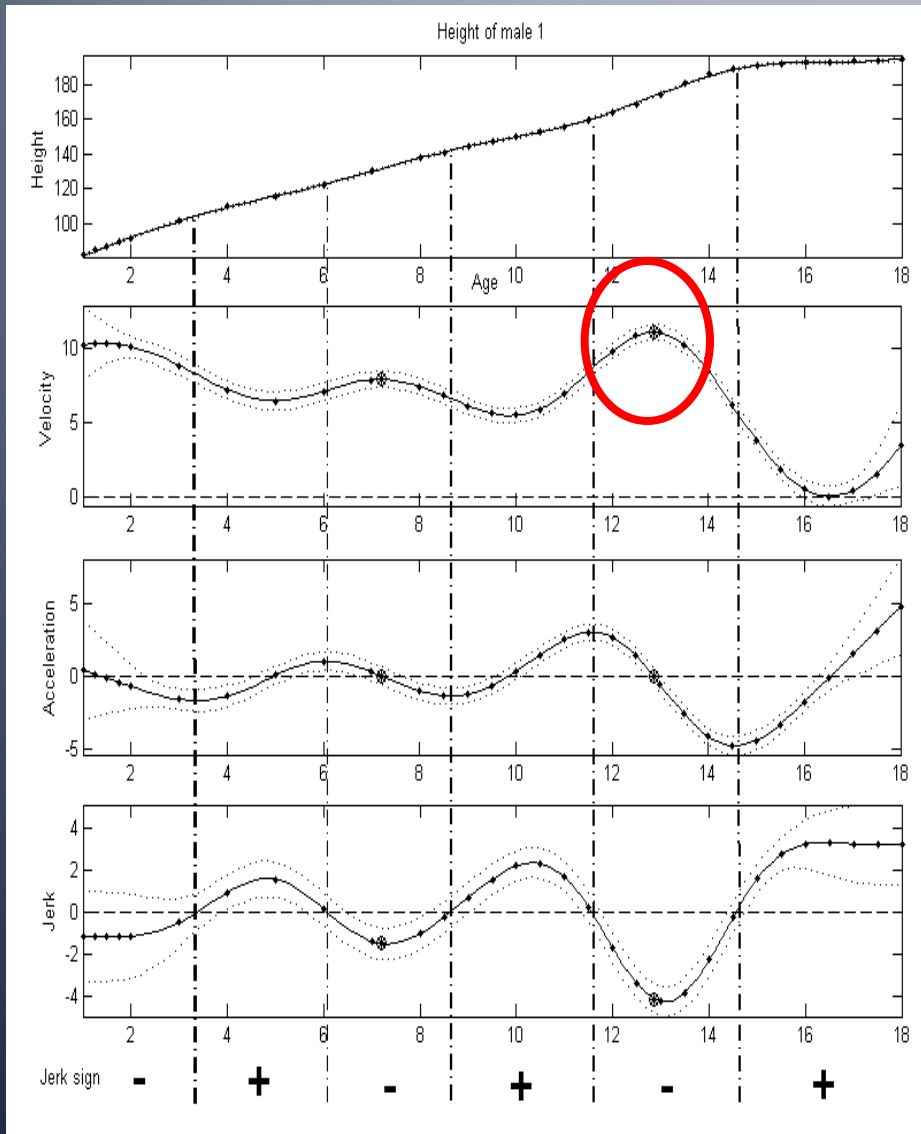
# Taille du vocabulaire

- Taille du lexique d'un jeune enfant



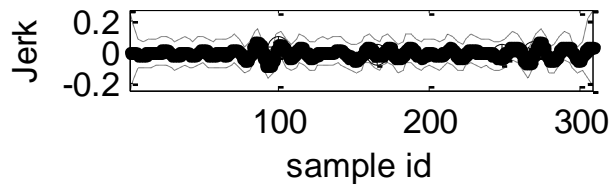
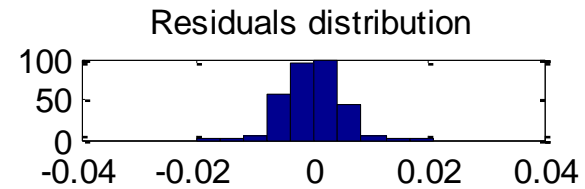
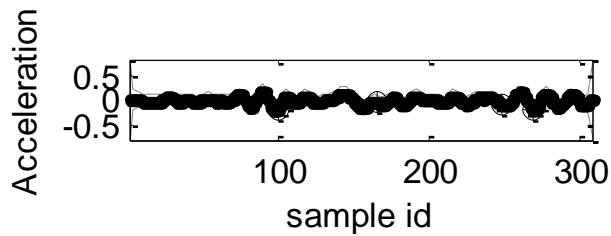
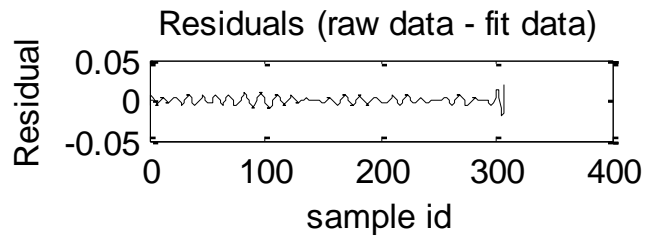
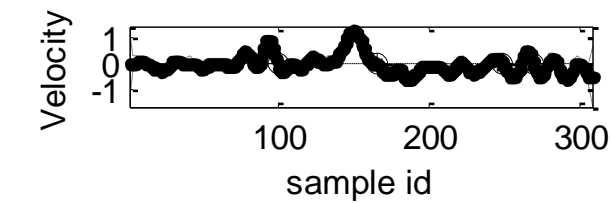
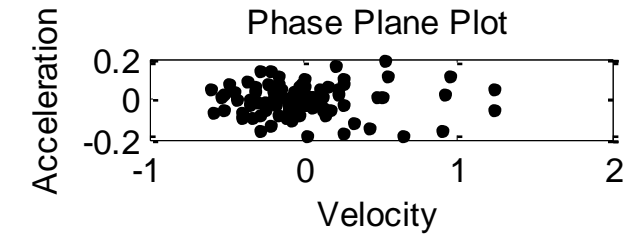
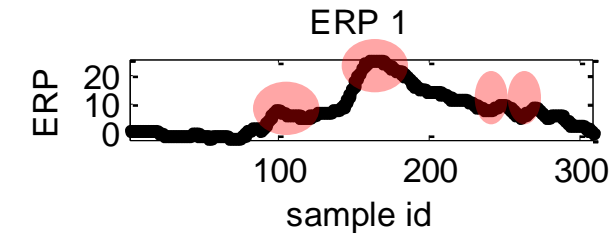
(Corrigan, 1978)

# Perspective: Données d'imagerie



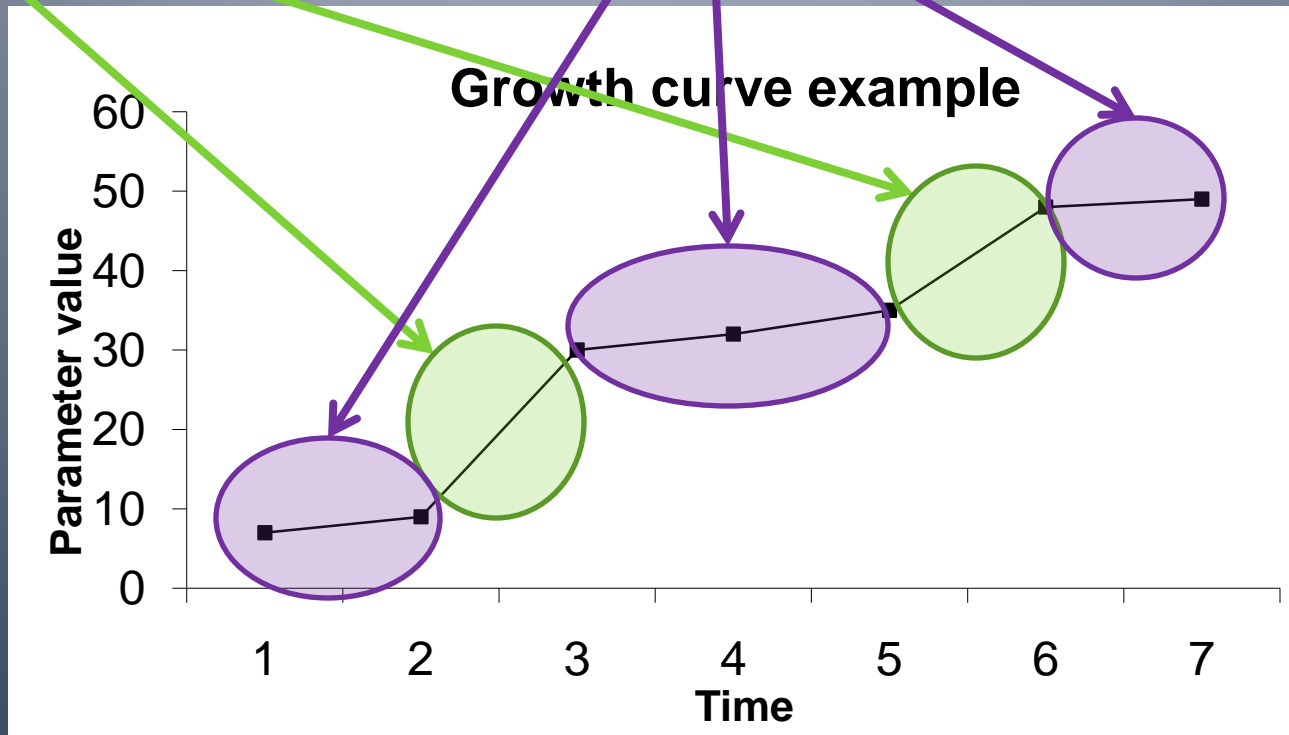
- AMD résout le problème général: trouver des maxima dans la fonction ou dans ses dérivées

# ERP



# Stades de développements

Poussées (transitions) et plateaux (stades) (Shultz, 2003)



# Sommaire

- AMD est un système automatisé qui augmente la rigueur de l'analyse des données longitudinales:
  - Critère statistique pour l'identification des maxima (poussées) significatifs
  - Quantification précise des maxima:
    - Début, Position, Amplitude & Durée
  - Outils de visualisation

# Remerciements

- Pour votre attention!
- Co-auteurs & collaborateurs
  - Thomas R. Shultz
  - James O. Ramsay
  - François Rivest
  - Signe Bray
  - Yoshio Takane
- Reference: Dandurand, F., & Shultz, T. R. (in press). Automatic detection and quantification of growth spurts. *Behavior Research Methods*.

# References

- Adolph, K. E., Robinson, S. R., Young, J. W., & Gill-Alvarez, F. (2008). What is the shape of developmental change? *Psychological Review*, 115(3), 527–543. doi: 10.1037/0033-295X.115.3.527
- Carlstein, E. G., Müller, H.-G., & Siegmund, D. (Eds.). (1994). *Change-point problems* (Vol. 23). Hayward, CA: Institute of Mathematical Statistics.
- Cleveland, W. S., Grosse, E., & Shyu, W. M. (1992). Local regression models. In J. M. Chambers & T. J. Hastie (Eds.), *Statistical Models in S* (309-376). New York: Chapman and Hall.
- Corrigan, R. (1978). Language development as related to stage 6 object permanence development. *Journal of Child Language*, 5, 173-189.
- Efron, B. (1981). Nonparametric estimates of standard error: The jackknife, the bootstrap and other methods. *Biometrika*, 68, 589-599.
- Efron, B. & Tibshirani, R. J. (1993). *An introduction to the bootstrap*. New York: Chapman & Hall.
- Eubank, R. L. (1999). *Spline smoothing and nonparametric regression*, second edition. New York: Marcel Dekker.
- Eveleth, P. B., & Tanner, J. M. (1990). *Worldwide variation in human growth*: Cambridge University Press.
- Flavell, J. H. (1971). Stage-related properties of cognitive development. *Cognitive Psychology*, 2, 421-453.
- Frisch, R. E., & Revelle, R. (1971). The height and weight of girls and boys at the time of initiation of the adolescent growth spurt in height and weight and the relationship to menarche. *Human Biology*, 43(1), 140-159.
- Ganger, J., & Brent, M. (2004). Reexamining the vocabulary spurt. *Developmental Psychology*, 40(4), 621-632.
- Green, P. J., & Silverman, B. W. (1994). *Nonparametric Regression and Generalized Linear Models*. London: CRC Press.
- Harrell, F.E. (2001). *Regression modeling strategies with applications to linear models, logistic regression and survival analysis*. New York: Springer-Verlag.
- Harezlak, J., & Heckman, N. E. (1999). *Crisp - a tool for bump hunting*. Vancouver: Department of Statistics, University of British Columbia.
- Hughes, B. D. (1996). *Random walks and random environments*. Oxford University Press.
- James, G. M. (2008). *Moments based functional synchronization*. Los Angeles, CA: University of Southern California.
- Jolicoeur, P., Pontier, J., & Abidi, H. (1992). Asymptotic models for the longitudinal growth of human stature *American journal of human biology* 4(4), 461-468.
- McMurray, B. (2007). Defusing the childhood vocabulary explosion. *Science*, 317(5838), 631.
- Mitchell, C. C., & McMurray, B. (2008). A stochastic model for the vocabulary explosion. In *Cognitive Science Conference Proceedings*. Washington, DC: Cognitive Science Society, Inc.
- Mitchell, C., & McMurray, B. (2009). On leveraged learning in lexical acquisition and its relationship to acceleration. *Cognitive Science*, 33(8), 1503-1523. doi: 10.1111/j.1551-6709.2009.01071.x
- Molinari, L., Largo, R. H., & Prader, A. (1980). Analysis of the growth spurt at age seven (mid-growth spurt). *Helv Paediatr Acta*, 35(4), 325-334.
- O’Sullivan, F. (1986). A statistical perspective on ill-posed inverse problems (c/r: P519-527). *Statistical Science* 1, 502–518.
- Piaget, J., & Inhelder, B. (2000). *The psychology of the child*. Basic Books.
- Ramsay, J. O., Bock, R. D., & Gasser, T. (1995). Comparison of height acceleration curves in the Fels, Zurich, and Berkeley growth data. *Annals of Human Biology*, 22(5), 413-426.
- Ramsay, J. O., Hooker, G., & Graves, S. (2009) *Functional Data Analysis with R and MATLAB*. New-York: Springer.
- Ramsay, J. O., & Silverman, B. W. (2005). *Functional data analysis*: New-York: Springer.
- Shawe-Taylor, J., & Cristianini, N. (2004). *Kernel Methods for Pattern Analysis*. Cambridge University Press.
- Shultz, T. R. (2003). *Computational developmental psychology*. Cambridge, MA: MIT Press.
- Siegler, R. S. (1998). *Children’s thinking*, 3rd edition. Upper Saddle River, NJ: Prentice-Hall Englewood Cliffs.
- Simonoff, J. S. (1996). *Smoothing methods in statistics*. New York: Springer.
- Tanner, J. M. (1978). *Growth at adolescence*, second edition: Blackwell Science Ltd.
- Tuddenham, R. D., & Snyder, M. M. (1954). *Physical growth of California boys and girls from birth to eighteen years* (Vol. 1): University of California Press.
- Van Geert, P. (1991). A dynamic systems model of cognitive and language growth. *Psychological Review*, 98 (1).
- Van Dijk, M., & Van Geert, P. (2007). Wobbles, humps and sudden jumps: A case study of continuity, discontinuity and variability in early language development. *Infant and Child Development*, 16, 7-33.