On the speed-accuracy trade-off and speed-curvature power law of tongue movements in repetitive speech

Stephan R. Kuberski, Adamantios I. Gafos



University of Potsdam Department of Linguistics and Excellence Area of Cognitive Science Potsdam, Germany







European Research Council Established by the European Commission

LabPhon16, June 19-22, 2018, Lisboa, Portugal

• Speed-curvature power law (also 1/3-power law)

- Speed-curvature power law (also 1/3-power law)
 - connects kinematic properties with geometrical characteristics

- **Speed-curvature power law** (also ½-power law)
 - connects kinematic properties with geometrical characteristics
- Fitts' law (spatially constrained speed-accuracy trade-off)

- **Speed-curvature power law** (also ½-power law)
 - connects kinematic properties with geometrical characteristics
- Fitts' law (spatially constrained speed-accuracy trade-off)
 - connects kinematic properties with task characteristics

- **Speed-curvature power law** (also ¹/₃-power law)
 - · connects kinematic properties with geometrical characteristics
- Fitts' law (spatially constrained speed-accuracy trade-off)
 - connects kinematic properties with task characteristics
- both laws are **effector-independent**: movements of hand, arm, legs, eyes, etc. (Hicheur et al., 2005; Plamondon and Alimi, 1997)

- **Speed-curvature power law** (also ½-power law)
 - · connects kinematic properties with geometrical characteristics
- Fitts' law (spatially constrained speed-accuracy trade-off)
 - connects kinematic properties with task characteristics
- both laws are **effector-independent**: movements of hand, arm, legs, eyes, etc. (Hicheur et al., 2005; Plamondon and Alimi, 1997)
- both laws also hold for perception of action (Viviani, 2002; Grosjean et al., 2007)

- **Speed-curvature power law** (also ¹/₃-power law)
 - · connects kinematic properties with geometrical characteristics
- Fitts' law (spatially constrained speed-accuracy trade-off)
 - connects kinematic properties with task characteristics
- both laws are **effector-independent**: movements of hand, arm, legs, eyes, etc. (Hicheur et al., 2005; Plamondon and Alimi, 1997)
- both laws also hold for perception of action (Viviani, 2002; Grosjean et al., 2007)

What are the implications for models of speech?

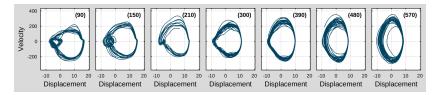
• paradigm of **repetitive speech** (cf. Kelso et al., 1985; Ostry et al., 1987; Patel et al., 1999)

- paradigm of **repetitive speech** (cf. Kelso et al., 1985; Ostry et al., 1987; Patel et al., 1999)
- systematic **speech rate control** by audible metronome
- speech rate from extremely slow to extremely fast (30-570 bpm)

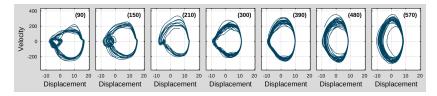
- paradigm of **repetitive speech** (cf. Kelso et al., 1985; Ostry et al., 1987; Patel et al., 1999)
- systematic **speech rate control** by audible metronome
- speech rate from extremely slow to extremely fast (30-570 bpm)
- sequences of syllables of the form /CV.../
- specifically: sequences of [tatata...] and [kakaka...]

- paradigm of **repetitive speech** (cf. Kelso et al., 1985; Ostry et al., 1987; Patel et al., 1999)
- systematic **speech rate control** by audible metronome
- speech rate from extremely slow to extremely fast (30–570 bpm)
- sequences of syllables of the form /CV.../
- specifically: sequences of [tatata...] and [kakaka...]
- about 4000 syllables of [ta] and [ka] each
- six native speakers of German and English

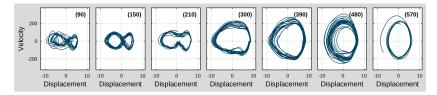
- **phase portraits** of tongue back state variables (displacement *x*, velocity \dot{x}) in sequences of **[taka...]**
- speech:



- phase portraits of tongue back state variables (displacement x, velocity x) in sequences of [taka...]
- speech:



• non-speech (trumpet double tonguing):



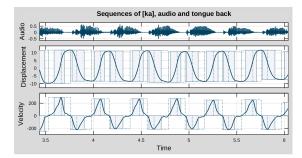
• electromagnetic articulometry (EMA, Carstens AG501)



- electromagnetic articulometry (EMA, Carstens AG501)
- high precision displacement data of selected articulators
- tongue tip for [t] and tongue back for [k] constrictions



- electromagnetic articulometry (EMA, Carstens AG501)
- high precision displacement data of selected articulators
- tongue tip for [t] and tongue back for [k] constrictions
- segmentation of continuous motion by zero velocity criterion





 standard model of speech gestures (linear oscillator with critical damping, Fowler et al., 1980; Browman and Goldstein, 1986; Saltzman and Munhall, 1989)

$$\ddot{x} = -kx - b\dot{x}, \quad b = 2\zeta\sqrt{k}, \quad \zeta = 1 \tag{1}$$

 standard model of speech gestures (linear oscillator with critical damping, Fowler et al., 1980; Browman and Goldstein, 1986; Saltzman and Munhall, 1989)

$$\ddot{x} = -kx - b\dot{x}, \quad b = 2\zeta\sqrt{k}, \quad \zeta = 1 \tag{1}$$

analytical solution reveals kinematic relations

$$v^* \propto \sqrt{k}A$$
 and $\frac{v^*}{A} \propto \frac{\pi}{T}$ (2)

 standard model of speech gestures (linear oscillator with critical damping, Fowler et al., 1980; Browman and Goldstein, 1986; Saltzman and Munhall, 1989)

$$\ddot{x} = -kx - b\dot{x}, \quad b = 2\zeta\sqrt{k}, \quad \zeta = 1 \tag{1}$$

analytical solution reveals kinematic relations

$$v^* \propto \sqrt{k}A$$
 and $\frac{v^*}{A} \propto \frac{\pi}{T}$ (2)

- direct proportionality of peak velocity v* and amplitude A
- inverse proportionality of v^*/A and duration T

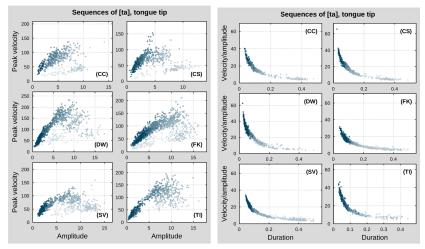
 standard model of speech gestures (linear oscillator with critical damping, Fowler et al., 1980; Browman and Goldstein, 1986; Saltzman and Munhall, 1989)

$$\ddot{x} = -kx - b\dot{x}, \quad b = 2\zeta\sqrt{k}, \quad \zeta = 1 \tag{1}$$

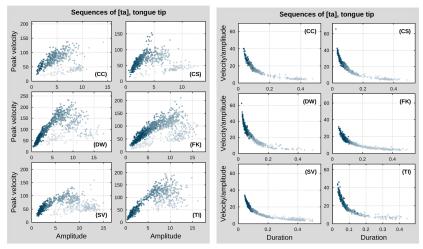
analytical solution reveals kinematic relations

$$v^* \propto \sqrt{k}A$$
 and $\frac{v^*}{A} \propto \frac{\pi}{T}$ (2)

- direct proportionality of peak velocity v^{*} and amplitude A
- inverse proportionality of v^*/A and duration T
- empirical evidence repeatedly reported, e.g., in Ostry et al. (1983), Munhall et al. (1985), and Vatikiotis-Bateson and Kelso (1990)



broad consistency with theoretical predictions and empirical reports



- · broad consistency with theoretical predictions and empirical reports
- divergence of A-v* correlation at larger amplitudes (slower rates) considered in Sorensen and Gafos (2016)

 relation between kinematics and geometry of movement known for a variety of motor actions (Hicheur et al., 2005)

- relation between kinematics and geometry of movement known for a variety of motor actions (Hicheur et al., 2005)
- the faster the speed of movement, the less curved its path (Viviani and Terzuolo, 1982)

- relation between kinematics and geometry of movement known for a variety of motor actions (Hicheur et al., 2005)
- the faster the speed of movement, the less curved its path (Viviani and Terzuolo, 1982)
- instantaneous speed ν is a **power function** of curvature κ (Lacquaniti et al., 1983)

$$\nu = k\kappa^{-\beta} \tag{3}$$

- relation between kinematics and geometry of movement known for a variety of motor actions (Hicheur et al., 2005)
- the faster the speed of movement, the less curved its path (Viviani and Terzuolo, 1982)
- instantaneous speed ν is a **power function** of curvature κ (Lacquaniti et al., 1983)

$$\nu = k\kappa^{-\beta} \tag{3}$$

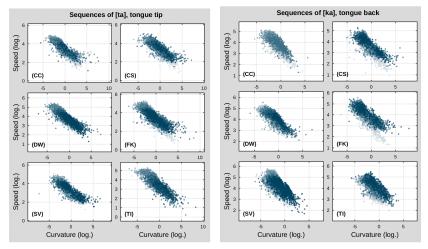
- velocity gain factor k segregates geometrical subunits
- exponent β is consistently found to be close to the value of $\frac{1}{3}$

• consider linear relation in log-log-transformed data

$$v = k\kappa^{-\beta} \implies \log v = \log k - \beta \log \kappa$$
 (4)

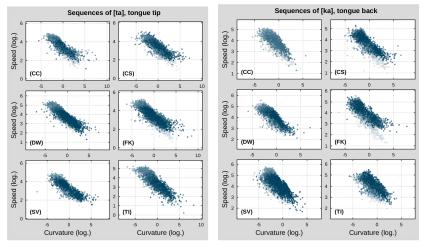
· consider linear relation in log-log-transformed data

$$v = k\kappa^{-\beta} \implies \log v = \log k - \beta \log \kappa$$
 (4)



· consider linear relation in log-log-transformed data

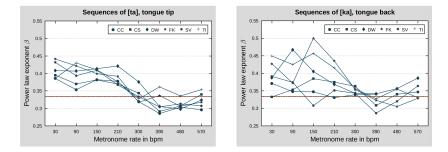
$$v = k\kappa^{-\beta} \implies \log v = \log k - \beta \log \kappa$$
 (4)



• clear evidence for the power law in speech

 consistent with the two earlier reports by Tasko and Westbury (2004) and Perrier and Fuchs (2008)

- consistent with the two earlier reports by Tasko and Westbury (2004) and Perrier and Fuchs (2008)
- new: significant **rate dependency** of the power law exponent β



 closest match with commonly found value of 1/3 at metronome rate of 300 bpm (fast speech rate)

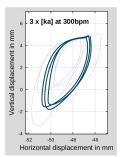
Why an exponent of 1/3?

Why an exponent of 1/3?

- movements can hardly be reduced to a single dimension (e.g., y-axis)
- traces form (at least) two-dimensional **spatial loops** (cf. Mooshammer et al., 1995; Birkholz et al., 2011)

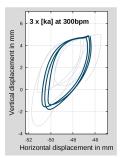
Why an exponent of 1/3?

- movements can hardly be reduced to a single dimension (e.g., y-axis)
- traces form (at least) two-dimensional **spatial loops** (cf. Mooshammer et al., 1995; Birkholz et al., 2011)
- our result: the faster the rate, the more elliptical the shape



Why an exponent of 1/3?

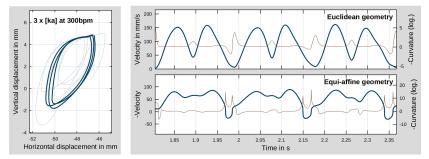
- movements can hardly be reduced to a single dimension (e.g., y-axis)
- traces form (at least) two-dimensional **spatial loops** (cf. Mooshammer et al., 1995; Birkholz et al., 2011)
- our result: the faster the rate, the more elliptical the shape



• Bennequin et al. (2009): **elliptical shapes** show analytically constant equi-affine speed with a power law exponent of **precisely** 1/3

Why an exponent of 1/3?

- movements can hardly be reduced to a single dimension (e.g., y-axis)
- traces form (at least) two-dimensional **spatial loops** (cf. Mooshammer et al., 1995; Birkholz et al., 2011)
- our result: the faster the rate, the more elliptical the shape



• Bennequin et al. (2009): elliptical shapes show analytically constant equi-affine speed with a power law exponent of precisely 1/3

• trade-off relation between **speed and accuracy** of aimed movement (Fitts, 1954; Fitts and Peterson, 1964)

- trade-off relation between **speed and accuracy** of aimed movement (Fitts, 1954; Fitts and Peterson, 1964)
- connects kinematics of movement (amplitude A and duration T) with task property (target extent W)

$$T = a + b ID, \quad ID = \log_2\left(\frac{A}{W} + 1\right) \tag{5}$$

- trade-off relation between **speed and accuracy** of aimed movement (Fitts, 1954; Fitts and Peterson, 1964)
- connects kinematics of movement (amplitude A and duration T) with task property (target extent W)

$$T = a + b ID, \quad ID = \log_2\left(\frac{A}{W} + 1\right) \tag{5}$$

• crucial requirements of Fitts' paradigm (Plamondon and Alimi, 1997):

- trade-off relation between **speed and accuracy** of aimed movement (Fitts, 1954; Fitts and Peterson, 1964)
- connects kinematics of movement (amplitude A and duration T) with task property (target extent W)

$$T = a + b ID, \quad ID = \log_2\left(\frac{A}{W} + 1\right) \tag{5}$$

- crucial requirements of Fitts' paradigm (Plamondon and Alimi, 1997):
 - 1. movements must be performed under temporal pressure

- trade-off relation between **speed and accuracy** of aimed movement (Fitts, 1954; Fitts and Peterson, 1964)
- connects kinematics of movement (amplitude A and duration T) with task property (target extent W)

$$T = a + b ID, \quad ID = \log_2\left(\frac{A}{W} + 1\right) \tag{5}$$

- crucial requirements of Fitts' paradigm (Plamondon and Alimi, 1997):
 - 1. movements must be performed under temporal pressure
 - 2. manipulate amplitude and target size (experimental conditions)

- trade-off relation between **speed and accuracy** of aimed movement (Fitts, 1954; Fitts and Peterson, 1964)
- connects kinematics of movement (amplitude A and duration T) with task property (target extent W)

$$T = a + b ID, \quad ID = \log_2\left(\frac{A}{W} + 1\right) \tag{5}$$

- crucial requirements of Fitts' paradigm (Plamondon and Alimi, 1997):
 - 1. movements must be performed under temporal pressure
 - 2. manipulate amplitude and target size (experimental conditions)

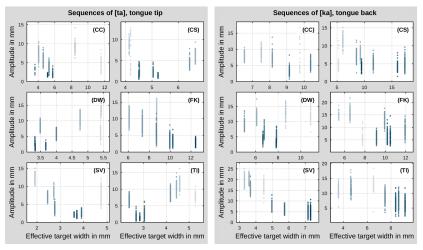
How to achieve this in speech?

• systematic **speech rate control** conforms closely to Fitts' paradigm:

systematic speech rate control conforms closely to Fitts' paradigm:
varying temporal pressure dependent on metronome rate

- systematic **speech rate control** conforms closely to Fitts' paradigm:
 - 1. varying temporal pressure dependent on metronome rate
 - 2. wide range of amplitudes and sufficient spread of widths

- systematic speech rate control conforms closely to Fitts' paradigm:
 - 1. varying temporal pressure dependent on metronome rate
 - 2. wide range of amplitudes and sufficient spread of widths



• computation of amplitudes *A* is straightforward: three-dimensional distance between **movement endpoints**

- computation of amplitudes *A* is straightforward: three-dimensional distance between **movement endpoints**
- endpoint distributions determine an *a posteriori* defined effective target width (Wobbrock et al., 2011)

$$W = \sqrt{2\pi e} \sigma, \quad \sigma = \sqrt{\frac{\sum_{i=1}^{N} \left\{ (x - \bar{x})^2 + (y - \bar{y})^2 + (z - \bar{z})^2 \right\}}{1 - N}} \quad (6)$$

- computation of amplitudes *A* is straightforward: three-dimensional distance between **movement endpoints**
- endpoint distributions determine an *a posteriori* defined effective target width (Wobbrock et al., 2011)

$$W = \sqrt{2\pi e} \sigma, \quad \sigma = \sqrt{\frac{\sum_{i=1}^{N} \left\{ (x - \bar{x})^2 + (y - \bar{y})^2 + (z - \bar{z})^2 \right\}}{1 - N}} \quad (6)$$

• both quantities A and W determine Fitts' index of difficulty

$$ID = \log_2\left(\frac{A}{W} + 1\right) \tag{7}$$

- computation of amplitudes A is straightforward: three-dimensional distance between movement endpoints
- endpoint distributions determine an *a posteriori* defined effective target width (Wobbrock et al., 2011)

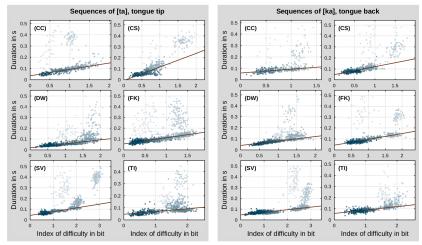
$$W = \sqrt{2\pi e} \sigma, \quad \sigma = \sqrt{\frac{\sum_{i=1}^{N} \left\{ (x - \bar{x})^2 + (y - \bar{y})^2 + (z - \bar{z})^2 \right\}}{1 - N}} \quad (6)$$

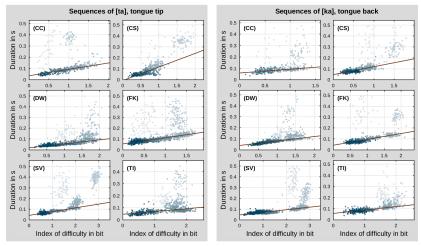
• both quantities A and W determine Fitts' index of difficulty

$$ID = \log_2\left(\frac{A}{W} + 1\right) \tag{7}$$

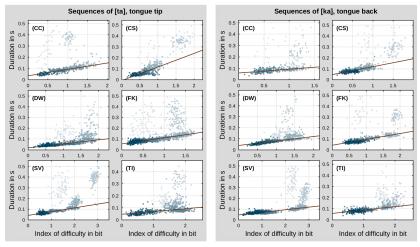
 observations of movement duration T should reveal the linearity of Fitts' law

$$T = a + b ID \tag{8}$$





• statistics on contiguous subsets of rates yields critical rates



- statistics on contiguous subsets of rates yields critical rates
- fast rates (≥150-210 bpm, dark shades) abide to Fitts' law
- slowest rates (<150-210 bpm, bright shades) do not

What are the implications for models of speech?

What are the implications for models of speech?

• Crossman and Goodeve (1983): Fitts' law holds true for any model implying **exponential decay** of distance to the target

What are the implications for models of speech?

- Crossman and Goodeve (1983): Fitts' law holds true for any model implying **exponential decay** of distance to the target
- all (non-trivial) solutions of the **standard model** possess an exponential signature

$$\ddot{x} = -kx - b\dot{x}, \quad b = 2\zeta\sqrt{k}, \quad \zeta > 0 \quad \Rightarrow \quad \text{Fitts' law}$$

What are the implications for models of speech?

- Crossman and Goodeve (1983): Fitts' law holds true for any model implying **exponential decay** of distance to the target
- all (non-trivial) solutions of the **standard model** possess an exponential signature

$$\ddot{x} = -kx - b\dot{x}, \quad b = 2\zeta\sqrt{k}, \quad \zeta > 0 \quad \Rightarrow \quad \mathsf{Fitts'} \ \mathsf{law}$$

But, what about the *absence* of Fitts' law at slower rates?

What are the implications for models of speech?

- Crossman and Goodeve (1983): Fitts' law holds true for any model implying **exponential decay** of distance to the target
- all (non-trivial) solutions of the **standard model** possess an exponential signature

$$\ddot{x} = -kx - b\dot{x}, \quad b = 2\zeta\sqrt{k}, \quad \zeta > 0 \quad \Rightarrow \quad \text{Fitts' law}$$

But, what about the *absence* of Fitts' law at slower rates?

 the model fails at those rates (<150–210 bpm) where one would expect it to apply

What are the implications for models of speech?

- Crossman and Goodeve (1983): Fitts' law holds true for any model implying **exponential decay** of distance to the target
- all (non-trivial) solutions of the **standard model** possess an exponential signature

$$\ddot{x} = -kx - b\dot{x}, \quad b = 2\zeta\sqrt{k}, \quad \zeta > 0 \quad \Rightarrow \quad \text{Fitts' law}$$

But, what about the *absence* of Fitts' law at slower rates?

- the model fails at those rates (<150–210 bpm) where one would expect it to apply
- other models with different dynamics exist but have not been investigated yetKröger et al., 1995; Sorensen and Gafos, 2016)

References

- Bennequin, Fuchs, Berthoz, and Flash (2009). "Movement timing and invariance arise from several geometries". PLoS Computational Biology, 5(7).
- Birkholz, Hoole, Kröger, and Neuschaefer-Rube (2011). "Tongue body loops in vowel sequences". Proceedings of the 9th International Seminar on Speech Production.

Browman and Goldstein (1986). "Towards an articulatory phonology". Phonology, 3(1).

- Crossman and Goodeve (1983). "Feedback Control of Hand-Movement and Fitts' Law". Quarterly Journal of Experimental Psychology, 35(2). Paper presented at the meeting of the Experimental Psychology Society, Oxford, 1963.
- Fitts (1954). "The information capacity of the human motor system in controlling the amplitude of movement". *Journal of Experimental Psychology*, 47(6).
- Fitts and Peterson (1964). "Information capacity of discrete motor responses". *Journal of Experimental Psychology*, 67(2).
- Fowler, Rubin, Remez, and Turvey (1980). "Implications for speech production of a general theory of action". Language Production: Speech and Talk. Ed. by B. Butterworth. New York, United States: Academic Press.
- Grosjean, Shiffrar, and Knoblich (2007). "Fitts's Law Holds for Action Perception". *Psychological Science*, 18(2).
- Hicheur, Vieilledent, Richardson, Flash, and Berthoz (2005). "Velocity and curvature in human locomotion along complex curved paths: a comparison with hand movements". *Experimental Brain Research*, 162(2).

References

- Kelso, Vatikiotis-Bateson, Saltzman, and Kay (1985). "A qualitative dynamic analysis of reiterant speech production: phase portraits, kinematics, and dynamic modeling". The Journal of the Acoustical Society of America, 77(1).
- Kröger, Schröder, and Opgen-Rhein (1995). "A gesture-based dynamic model describing articulatory movement data". The Journal of the Acoustical Society of America, 98(4).
- Lacquaniti, Terzuolo, and Viviani (1983). "The law relating the kinematic and figural aspects of drawing movements". Acta Psychologica, 54(1–3).
- Mooshammer, Hoole, and Kühnert (1995). "On loops". Journal of Phonetics, 23(1-2).
- Munhall, Ostry, and Parush (1985). "Characteristics of velocity profiles of speech movements". Journal of Experimental Psychology: Human Perception and Performance, 11(4).
- Ostry, Keller, and Parush (1983). "Similarities in the control of the speech articulators and the limbs: kinematics of tongue dorsum movement in speech". *Journal of Experimental Psychology: Human Perception and Performance*, 9(4).
- Ostry, Cooke, and Munhall (1987). "Velocity curves of human arm and speech movements". Experimental Brain Research, 68(1).
- Patel, Löfqvist, and Naito (1999). "The acoustics and kinematics of regularly timed speech: a database and method for the study of the p-center problem". Proceedings of the 14th International Congress of Phonetic Sciences. Vol. 1.
- Perrier and Fuchs (2008). "Speed-curvature relations in speech production challenge the 1/3 power law". *Journal of Neurophysiology*, 100(3).

References

- Plamondon and Alimi (1997). "Speed/accuracy trade-offs in target-directed movements". Behavioral and Brain Sciences, 20(2).
- Saltzman and Munhall (1989). "A dynamical approach to gestural patterning in speech production". *Ecological Psychology*, 1(4).
- Sorensen and Gafos (2016). "The gesture as an autonomous nonlinear dynamical system". *Ecological Psychology*, 28(4).
- Tasko and Westbury (2004). "Speed-curvature relations for speech-related articulatory movement". *Journal of Phonetics*, 32(1).
- Vatikiotis-Bateson and Kelso (1990). "Linguistic structure and articulatory dynamics: a cross language study". Haskins Laboratories Status Report in Speech Research, SR-103/104.
- Viviani (2002). "Motor competence in the perception of dynamic events: a tutorial". Common mechanisms in perception and action: attention and performance XIX. ed. by W. Prinz and B. Hommel. Oxford, United Kingdom: Oxford University Press.
- Viviani and Terzuolo (1982). "Trajectory determines movement dynamics". Neuroscience, 7(2).
 Wobbrock, Shinohara, and Jansen (2011). "The effects of task dimensionality, endpoint deviation, throughput calculation, and experiment design on pointing measures and models". Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.

Thank you very much.





European Research Council Established by the European Commission

Support from DFG SFB 1287 and ERC AdG 249440 is gratefully acknowledged.



The Collaborative Research Centre SFB 1287 "Limits of Variability in Language: Cognitive, Grammatical, and Social Aspects" in Potsdam, Germany, invites applications for short-term fellowships available in 2019, 2020.

We are offering several three- to twelve-month fellowships to external PhD candidates who would like to do their research within one dedicated project of the CRC. Please see the available projects here:

https://www.uni-potsdam.de/en/sfb1287/subprojects.html

The monthly fellowship is tax exempt and covers a basic amount (1365 EUR) plus direct costs (103 EUR) plus 400 EUR per child, if applicable. Holders of the grant need to cover health insurance on their own. See

https://www.uni-potsdam.de/sfb1287/index.html