



Translating from Unisemiotic to Polysemiotic Narratives:

A study of Finnish speech and gestures

Karoliina Louhema, Jordan Zlatev, Maria Graziano & Joost van de Weijer
Lund University



How does the semiotic system of a source narrative influence how the story is re-narrated?

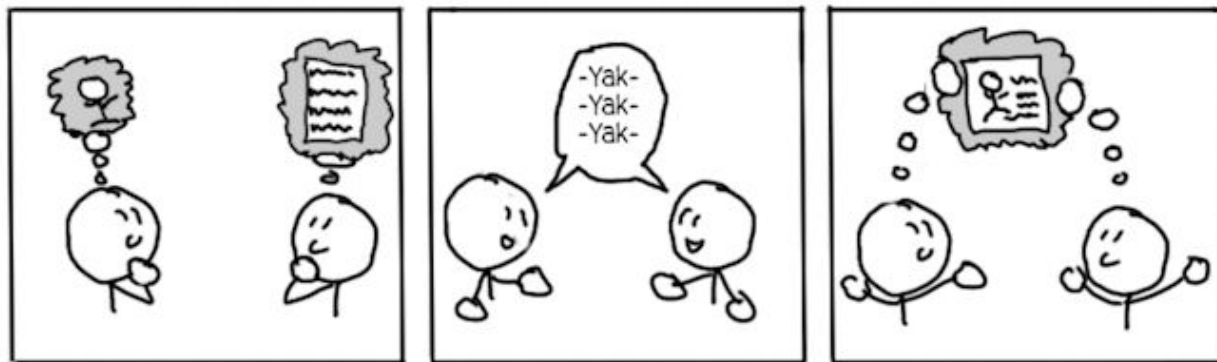
If a story is experienced in:

- **pictures** (depiction only condition)
- **speech** (language only condition)

...does it result in differences in the retelling of that story? Does it lead to different *polysemiotic narratives* (speech + gestures), where some characteristics of the source narrative system (depiction or language) is reflected in the retelling performance?

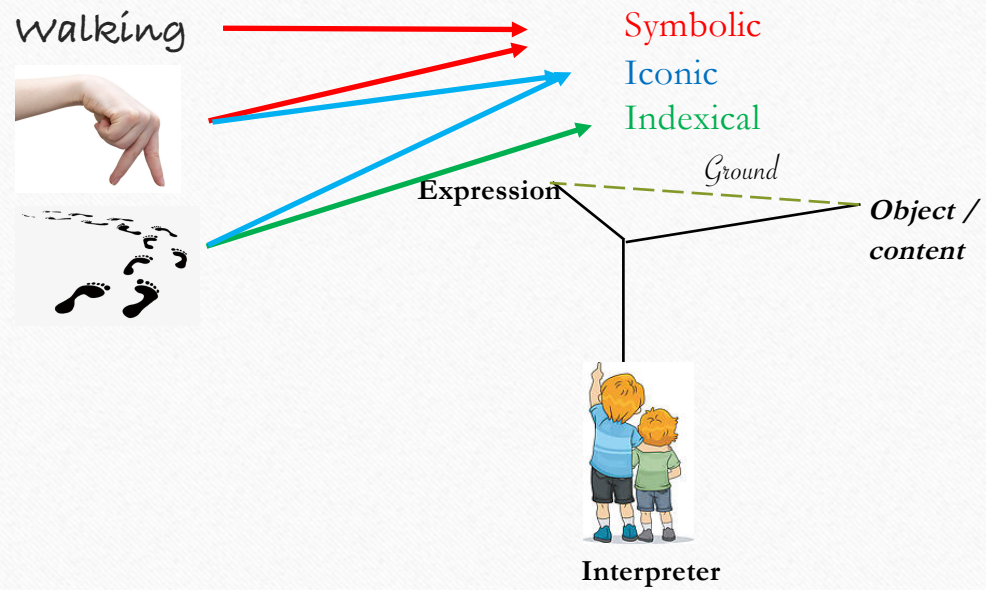
Semiotic systems: language, gestures, depiction
Sensory modalities: vision, hearing, touch

Polysemiotic? Multimodal?



Sign systems

(e.g. Sonesson, 2010; Ahlner & Zlatev, 2010).



Properties	Semiotic systems		
	Language (Speech, Writing, Signing)	Gesture	Depiction
Production	Vocal, Material, Body	Body	Material
Perception (dominant)	Auditory, Visual, Visual	Visual	Visual
Rapid fading	Yes, No, Yes	Yes	No
Double articulation	Yes, Yes, Yes?	No	No
Semiotic grounds	Conventional > Iconic, Indexical	Iconic/Indexical/Convent ional	Iconic > Indexical, Conventional
Syntagmatic relations	Compositional	Sequential	Possibly sequential

(Zlatev, 2019; Stampoulidis et al., in press)

Ideophones



Unimodal (onomatopoeia)

=sound to sound resemblance

Splash!



Crossmodal

Words that phonologically “depict” something in some other sensory modality than audition – often combining different modalities.

(Dingemanse, 2018).

Gestures

(Kendon, 2004; Zlatev, 2015)

1. Iconic

(Icon – similarity based sign)



- Enacting gestures
- Symbolic gestures: (Zlatev, 2015)
- drawing, molding, representing (Müller, 2014)

2. Deictic

(Index – contiguity-based sign)



4. Pragmatic

- multifunctional
- e.g. structure and regulate interaction

3. Emblematic

(Symbol – convention-based sign)



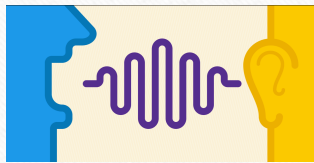
Narratives: representations of sequences of events (e.g. Prince 1982)

- **Speech narrative** (audio modality)
 - Temporally successive, rapid fading → logic of time
 - Prosodic variation to highlight and frame
- **Picture narrative** (visual modality)
 - Framed space → logic of space
 - Colour to highlight and frame
- **Polysemiotic** (multimodal) **narrative** (= speech + gestures)
 - Logic of time + logic of space

Intersemiotic translation

What happens when a story expressed in either *speech* or *pictures* is translated into a *polysemiotic narrative* (speech + gestures)?

- Would HEARING the story give more *coherent* retellings of the story in respect to:



Coherence

- Organisation
- Development of the plot

“...One night while the boy and his dog were sleeping, the frog climbed out of the jar. He jumped out of an open window. When the boy and the dog woke up the next morning...”

- Would SEEING the story give rise to more *perceptually detailed* narratives, including :



Iconicity

- Iconic gestures (esp. enactments)



Design of the experiment

- 38 native Finnish speakers in two groups were first presented the **same story**, *Frog, Where are You?* (M. Mayer, 1969), either in *pictures* or *language* (speech), which they then re-narrated in speech (and gesture) to an interlocutor.
 - **picture narrative**: 24 still pictures that each represent one or more events, and together the pictures constitute a story (the original black-and-white pictorial narrative was added **colour** so that the characters and events became more easily foregrounded, in a similar manner that **prosodic variation** is used in speech to foreground and emphasise)
 - **speech narrative**: an oral version of the same story in Finnish

Procedure

- Both groups were exposed to the source narratives only once
- The length of the audio narrative was 3,5 minutes, and in order to make the length of the narratives identical between the groups, the sequence of pictures was shown in the same pace as the audio recording

Analysis

- The reproduced narratives were video-recorded, and later transcribed and annotated in detail using multimedia annotator ELAN 4.9.1-b
- Speech was analysed for clauses, words, plot elements, and connective devices
- Gestures were analysed for their main function: iconic (enacting and symbolic), deictic, emblematic, pragmatic
- Statistical results based on Poisson regression (for rates) and logistic regression (for proportions). The main predictor was always semiotic system

Hypotheses

1. **There will be a higher number of plot elements when translated from speech narratives** due to linguistic cohesion in the source narrative, which provides the foundation for the (linear) unfolding of the narrative plot.
2. **There will be a more diverse use of connective devices (in terms of function) when translated from speech** due to their presence in the source narrative, whereas in the picture source narrative the relations between successive events need to be inferred.
3. **There will be a higher number of iconic gestures when translating from picture narrative**, reflecting the more iconic nature of the source narrative.
4. **There will be more enacting gestures when translating from pictures** compared to speech narratives, as these are the kind that most closely correspond to the primary iconicity of pictures in the source narrative.

Source = Depiction condition



Source = Language condition



	Total		Proportions		Rate per clause (per word)	
	P	S	P	S	P	S
Narratives	19	19				
<i>Words</i>	7848	6822				
<i>Clauses</i>	1285	1117				
<i>Average clause length (words)</i>	6.10	6.10				
H1 Narrative structure / plot elements (clauses)	1237	1042	0.97	0.93		
<i>introduction</i>	136	61	0.11	0.06		
<i>onset of the plot</i>	86	72	0.07	0.07		
<i>unfolding of the plot</i>	895	763	0.72	0.73		
<i>resolution of the plot</i>	109	141	0.09	0.14		
<i>ending</i>	11	5	0.01	0.00		
Metanarrative clauses	42	66	0.03	0.06		
Extranarrative clauses	6	9	0.00	0.01		
H2 Connective devices	931	719			0.69	0.75
<i>additive</i>	433	318	0.46	0.44		
<i>adversative</i>	42	75	0.05	0.10		
<i>causal</i>	84	46	0.09	0.06		
<i>temporal</i>	295	253	0.32	0.35		
<i>continuatives</i>	77	27	0.08	0.04		

Results: speech

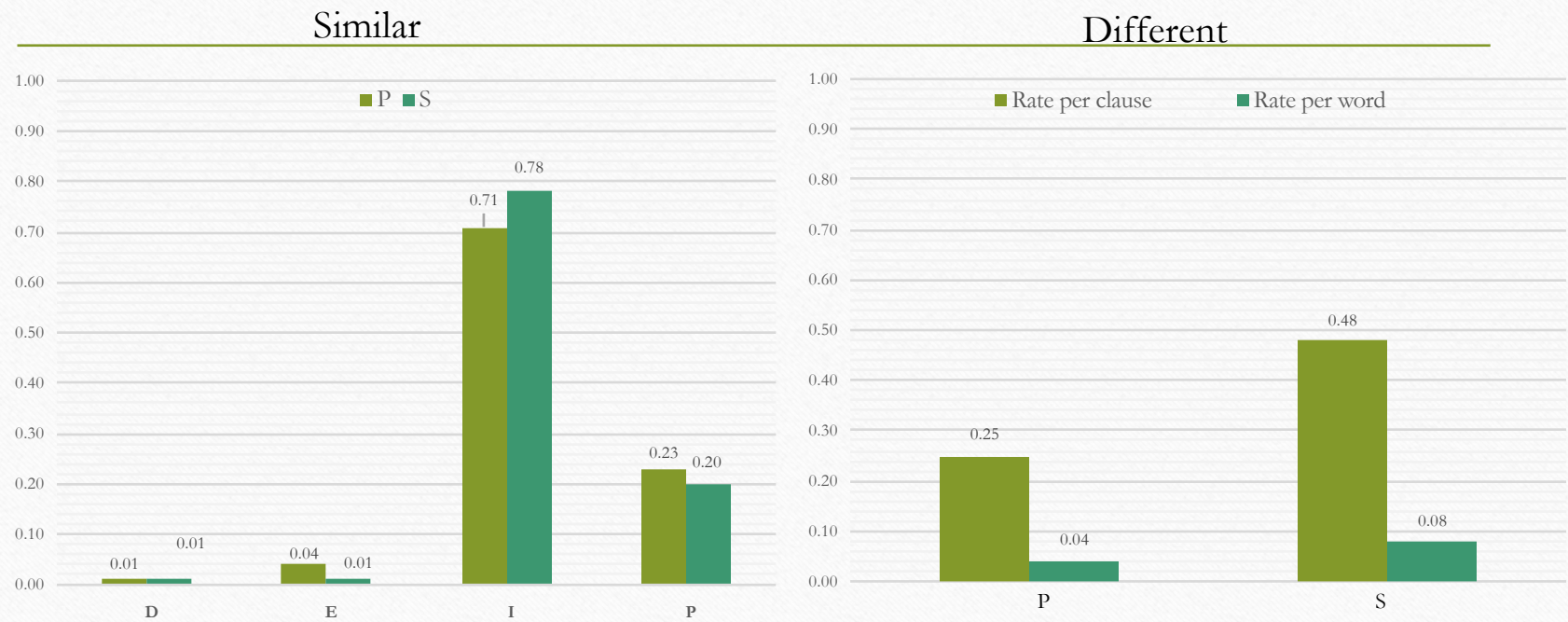
- Narratives in picture condition were longer than those in speech condition
- More non-narrative clauses in speech condition
- Significant difference in two types of connective devices (in terms of function) between the groups may be connected to the source narrative systems

Results: gestures

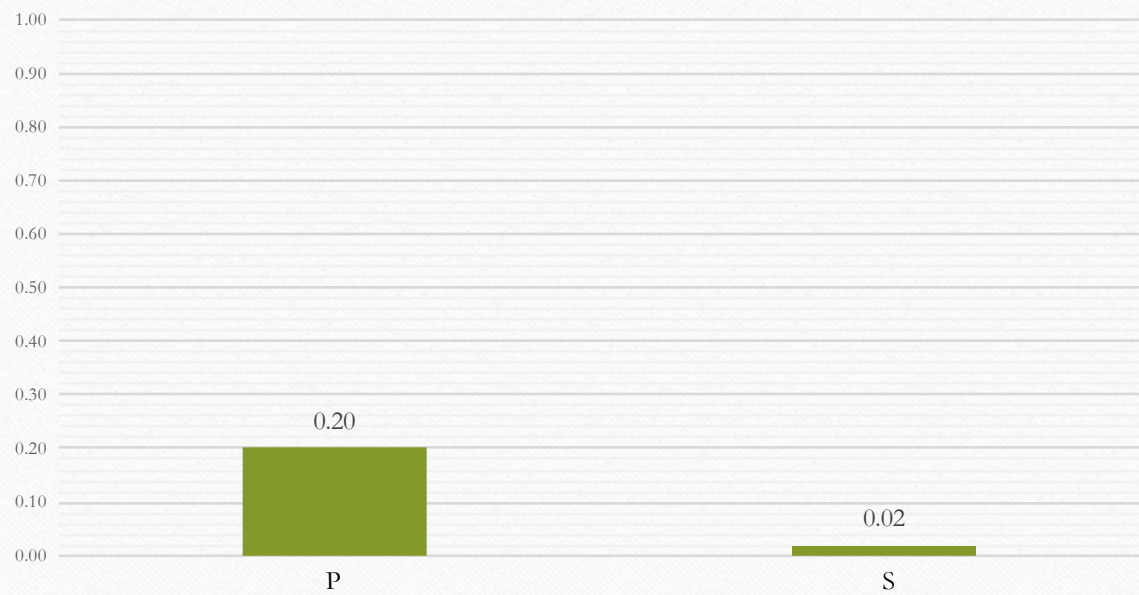
	Total		Proportions		Rate per clause (per word)	
	P	S	P	S	P	S
H3 – H4 Gestures	321	534			0.259 (0.04)	0.512 (0.08)
<i>deictic</i>	3	6	0.01	0.01		
<i>emblematic</i>	14	4	0.04	0.01		
<i>iconic</i>	229	418	0.71	0.78		
- <i>enacting gestures</i>	46	10	0.20	0.02		
- <i>non-enacting gestures</i>	183	408	0.80	0.98		
<i>pragmatic</i>	75	106	0.23	0.20		

- Overall more gestures in speech condition than in picture condition
→ **Not expected!**
- Also more Iconic gestures in narratives translated from speech than in the picture condition
→ **Not expected**
- But, significantly more enacting gestures in picture condition than in speech condition
→ **As expected!**

Gesture proportions



Enacting gestures



Discussion

Although three of the four hypotheses were not supported, many differences were found between the conditions that made the two retellings of the *Frog story* different:

- Narratives in **speech condition** had significantly more **adversative connectives** (“but”), as in:

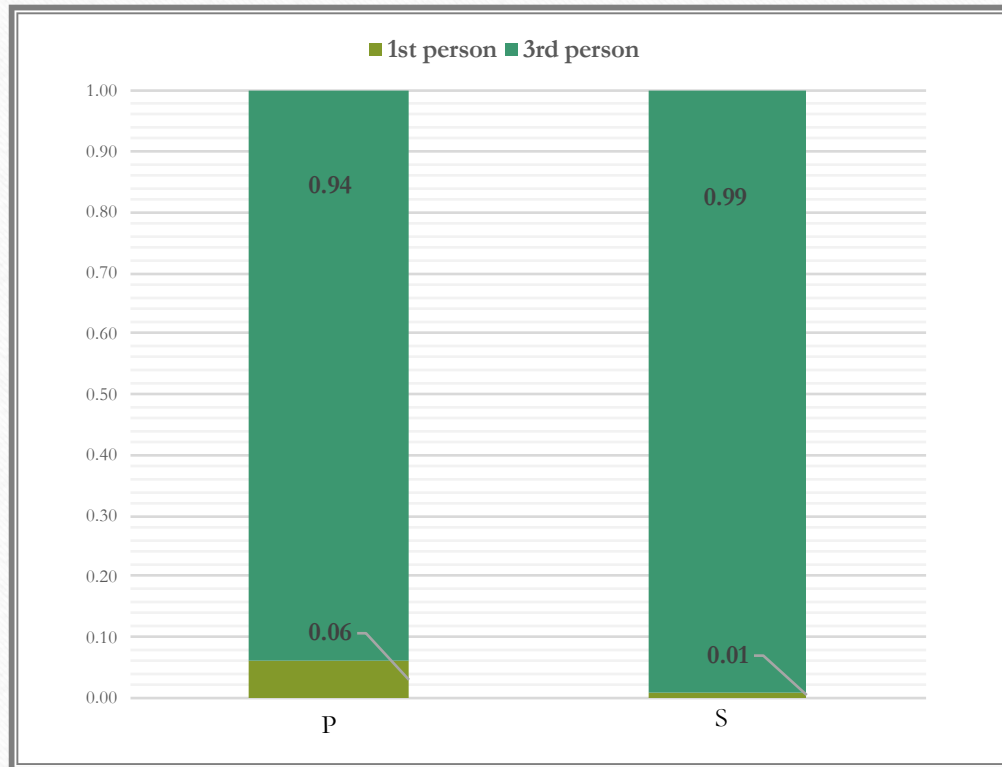
sitä tarina ei kertonut . mutta poika oli onnellinen .

‘the story didn’t tell that . **but** the boy was happy .’

- Narratives in **picture condition** had significantly more **continuative connectives**, such as “anyway”, “well”, “that”

että älä mee sinne . et ampiaiset tulee . ja pistää sinua .

‘that don’t go there . **that** the bees will come . and sting you .’



Proportion of first person direct speech and third person indirect speech perspective.

- More first person perspective in narration (direct speech) in picture condition

*ja mennään kysymään metsän eläimiltä vinkejä .
että hei että ootko nähny nähny missään sammakkoa?*

'and let's go and ask the forest animals for some tips . **that** hey have you seen the frog somewhere?'

- More enacting gestures in picture condition
- Although the narratives in speech condition were shorter, they still took longer to narrate (time), and also the overall number of gestures was higher

Experientiality

(Fludernik, 1996)

! The factor that could differentiate the source language and source depiction conditions from one another

- The consciousness of the protagonist or the consciousness of the narrator is made manifest
- Experiential value expressed through empathy or “perceptual focalisation” (i.e. access to the consciousness of the character)

👉 The level of *experientiality* appeared to be higher in source = depiction condition

- More “experiential value” through empathy:

ja eihän Max kun ei ollu aikasemmin nähny . ni ei tienny . et se oli ampiaispesä .

‘and of course since Max had never seen alike before . he didn’t know . that it was a beehive .’

- More “experiential value” through perceptual focalisation

Source = Depiction condition

— Narratives seemed to have been internalised better

→ more creative, vivid (enacting gestures and direct speech),
narratives “added” to the story (they were e.g. personalised by naming
the characters, by various introductions to the introduction, and giving
motivations to actions)

👉 Pictures offer a greater **semiotic freedom of interpretation** compared to language

Source = Language condition

- The narratives seemed not as well internalised
 - appeared less “fluent”, and contained more hesitation (affecting the time used for narration) and non-narrative clauses like:

ja en muista miten se liittyy tähän tarinaan
'and I can't remember how it's related to this story'

- 👉 Narrative style more focused on accuracy than empathy
 - Less semiotic freedom of interpretation
(who did what to whom? In which order? In what location?)

Depiction vs. Language

- Translations from picture narratives were more comparable to genuine *storytelling* and more often took the *inside-the-storyworld* perspective
 - 1st person perspective in narration (direct speech) and gestures (enactments)
- Translations from speech narratives were closer to reports of the events narrated from the *outside-the-storyworld* perspective
 - 3rd person perspective in narration in both speech (indirect speech) and gestures

Outside vs. inside the storyworld

- What could be the reason for the substantially different rates of gestures between speech and picture condition?



Managing the visual map of the storyworld from the outside would require a greater and more systematic use of gestures than when one positions oneself inside the story, being able to refer to things in relation to one's own body.

- Reference tracking is a bimodal and polysemiotic
 - Gesture space can be used effectively in order to create cohesion

(e.g. Kendon, 2004)

Gestures in target narratives

- Gesturing may help shifting load from verbal working memory to other cognitive systems or external representations (Goldin-Meadow, Nusbaum, Kelly & Wagner, 2001)
 - a facilitative role of gestures on narrative production?
(esp. iconic “representing” gestures?)
- Iconic enacting gestures (and those that depict by molding and drawing?) are likely to be elicited when visual, motoric, or spatial information is translated into speech (Rimé and Schiaratura, 1991)
 - a way to make narratives more vivid (requires a high level of experientiality)

Conclusions

The source narrative system may affect how a story is experienced and how polysemiotic retellings of the story are constructed!

- Internalisation
- Experientiality
- Perspective (1st person/3rd person)
- Gesture rate
- Connective devices
- Vividness and “fluidity”
- ...

Thank you! 😊



References

-
- Ahlner, F., & Zlatev, J. (2010). Cross-modal iconicity: A cognitive semiotic approach to sound symbolism. *Sign Systems Studies* 38(1/4), 298-348.
 - Dingemanse, M. (2018). Redrawing the margins of language: Lessons from research on ideophones. *Glossa: A Journal of General Linguistics* 3(1): 4, 1-30. DOI: <https://doi.org/10.5334/gjgl.444>
 - Fludernik, M. (1996). *Towards a 'natural' narratology*. New York: Routledge.
 - Goldin-Meadow, S., Nusbaum, H., Kelly, S. D., & Wagner, S. (2001). Explaining math: Gesturing lightens the load. *Psychological Science* 12(6), 516-522.
 - Kendon, A. (2004). *Gesture: Visible action as utterance*. New York: Cambridge University Press.
 - Mayer, M. (1969). *Frog, where are you?* New York: Dial Press.
 - Müller, C. (2014). Gestural modes of representation as techniques of depiction. In *Body – Language – Communication: An international handbook on multimodality in human interaction 2*, 1687-1702. Berlin: De Gruyter Mouton.
 - Prince, G. (1982). *Narratology: The Form and Functioning of Narrative*. Mouton.
 - Zlatev, J. (2015b). The emergence of gestures. In B. MacWhinney & W. O'Grady (Eds.) *The Handbook of language emergence*. Malden, MA.: Wiley-Blackwell Publishing.
 - Zlatev, J. (2019)

References:

- Image polysemiotic com: "Communication foundations" (2013). Lynda, LinkedIn. Retrieved from: <https://cdn.lynda.com/course/141501/141501-636165434799010935-16x9.jpg>.
- Image speech - hearing: Biswas, J. (2018). "Behind Tacotron 2: Google's Incredibly Real Text To Speech System". *Analytics India Magazine*. Retrieved from: https://analyticsindiamag.com/wp-content/uploads/2018/01/201504_blog_slp_v1.png
- Image iconic gesture: "3: I don't know". In "Five UAE hand gestures that you need to know". *Whatson.ae*. (2018). Retrieved from: <http://whatson.ae/dubai/wp-content/uploads/2016/10/asdasd.jpg>
- Image (signlang): "Augmented Reality App Created by NYU Students Can Translate Sign Language in Real Time" (2018). *Communications ACM*. Retrieved from: https://cacm.acm.org/system/assets/0003/0765/040318_cuchimes.com_ASL.large.jpg?1522773952&1522773951
- Image multimodal: "3 Resources to Help You Create Remarkable Visual Content" (2015). *Copyblogger*. Retrieved from: <https://www.copyblogger.com/images/2014/12/image-brainstorming-process.jpg>
- Image footprints: "Walking footprint". Pngtree. Retrieved from: https://png.pngtree.com/element_origin_min_pic/17/07/07/6355530a728edceb3835827e9c98f30e.jpg
- Image looking: "Clip Art of Back View of Boy Siblings Looking Up". *Photosearch*. from: http://photos2.fotosearch.com/bthumb/CSP/CSP992/back-view-of-boy-siblings-looking-up-clip-art_k14570286.jpg
- Image walking: "Study reveals surprising health benefits of walking barefoot (2018). *Jacarandajm*. Retrieved from: https://turntable.kagiso.io/images/walking_barefoot_istock.width-800.jpg
- Image splash: "The research that definitely isn't making a splash" (2016). *Oxford Science blog*. Retrieved from: http://www.ox.ac.uk/sites/files/oxford/styles/ow_medium_feature/public/field/field_image_main/splash.jpg?itok=WMRCl-rT.
- Image meow: "Cute-kitty-animated-gif-25.gif". *The Transformice Cafe Wiki*. Retrieved from: <https://encrypted-tbn0.gstatic.com/images?q=tbn:ANd9GcT3vNXAZfwT1TDpCK0WwEj-SLdEaLhuvi29AOp3QVE8PDgi97ns>
- Image peace: Pngtree. Retrieved from: https://png.pngtree.com/element_origin_min_pic/17/02/11/4e26670bff2f0cbd07eae62058bec816.jpg
- Image deictic gesture: "Gesture Use between Parents and Children with Autism" (2014). Sydney Cognitive Development Centre. Retrieved from: <https://scdcentre.com/wp-content/uploads/2014/09/Autism-Gestures.jpg>