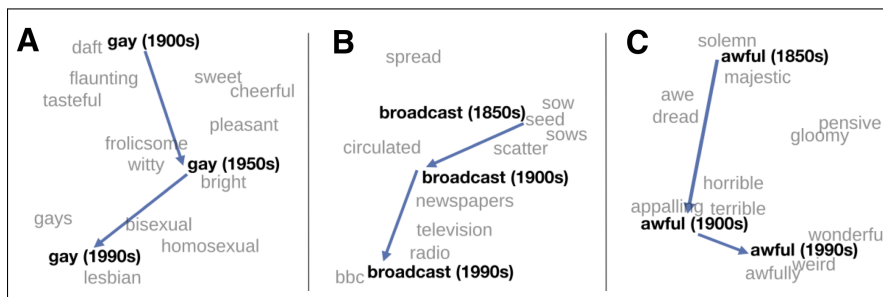


Figure 2  
Sample subtrees from a 1,000-word mutual information tree.

FIGURE K.8 – Dendrogramme illustrant le résultat d’un clustering hiérarchique, tiré de Brown *et al.* (1992), voir aussi (Smith, 2020)

- (5) a. Pen : pencil, ink, paper, write
- b. Pen : pencil, fountain, ink, paper, write
  
- (6) a. Spots : dog, dirty, dirt, stripes, dark
- b. Spots : spot, found, giraffe, bald, covered



**Figure 6.14** A t-SNE visualization of the semantic change of 3 words in English using word2vec vectors. The modern sense of each word, and the grey context words, are computed from the most recent (modern) time-point embedding space. Earlier points are computed from earlier historical embedding spaces. The visualizations show the changes in the word *gay* from meanings related to “cheerful” or “frolicsome” to referring to homosexuality, the development of the modern “transmission” sense of *broadcast* from its original sense of sowing seeds, and the pejoration of the word *awful* as it shifted from meaning “full of awe” to meaning “terrible or appalling” (Hamilton et al., 2016).

Both		Topics only		LSA only		Neither	
Cue		Cue		Cue		Cue	
CURB	UNCOMMON	PEN	SKILLET	DESTRUCTION	SEPARATE	SPOTS	RICE
Associates		Associates		Associates		Associates	
STREET	COMMON	PENCIL	PAN	DESTROY	DIVIDE	DOG	CHINESE
SIDEWALK	RARE	INK	FRY	WAR	DIVORCE	DIRTY	WEDDING
ROAD	WEIRD	PAPER	EGG	RUIN	PART	DIRTY	FOOD
CAR	UNUSUAL	WRITE	COOK	DEATH	SPLIT	STRIPES	WHITE
TIRE	UNIQUE		IRON	KILL	REMOVE	DARK	CHINA
LSA		LSA		LSA		LSA	
STREET	COMMON	HOG	COOKING	DESTROY	DIVIDE	SPOT	PADDY
PEDESTRIAN	FREQUENT	HEN	COOKED	VULNERABLE	INDEPENDENT	GIRAFFE	HARVEST
TRAFFIC	CIRCUMSTANCE	NAP	OVEN	BUMBLE	MIXTURE	GRAY	WHEAT
SIDEWALK	COUPLE	FIX	FRIED	THREAT	ACCOUNT	HIKE	BARLEY
AVENUE	WHALE	MOP	COOK	BOMB	COMMA	MILDEW	BEANS
(1)	(1)	(7)	(6)	(1)	(1)	(2972)	(322)
Topics		Topics		Topics		Topics	
STREET	COMMON	PENCIL	PAN	WAR	FORM	SPOT	VILLAGE
CAR	CASE	FOUNTAIN	KITCHEN	BUCK	SINGLE	FOUND	CORN
CORNER	BIT	INK	COOKING	TAP	DIVISION	GIRAFFE	WHEAT
WALK	EASY	PAPER	STOVE	NUCLEAR	COMMON	BALD	GRAIN
SIDEWALK	KNOWLEDGE	WRITE	POT	DAMAGE	DIVIDE	COVERED	FOOD
(1)	(1)	(1)	(1)	(24)	(5)	(1563)	(68)

Figure 9. Actual and predicted associates for a subset of cues. Two cues were randomly selected from the sets of cues for which (from left to right) both models correctly predicted the first associate, only the topic model made the correct prediction, only latent semantic analysis (LSA) made the correct prediction, and neither model made the correct prediction. Each column lists the cue, human associates, predictions of the topic model, and predictions of LSA, presenting the first five words in order. The rank of the first associate is given in parentheses below the predictions of the topic model and LSA.

FIGURE K.9 – Table extraite de (Griffiths *et al.*, 2007) avec des exemples de comparaison entre associations humaines et associations produites par des modèles distributionnels (LSA/topic modelling) de première génération

CUE	RESP	#G	#P	FSG	BSG
lunch	dinner	156	42	0.269	0.096
lunch	food	156	32	0.205	0.011
lunch	eat	156	13	0.083	0.0
lunch	meal	156	10	0.064	0.063
lunch	box	156	9	0.058	0.0
lunch	sandwich	156	9	0.058	0.037
lunch	noon	156	6	0.038	0.200
noon	lunch	150	30	0.200	0.038
noon	twelve	150	22	0.147	0.034
noon	sunshine	150	20	0.133	0.0
food	eat	180	73	0.406	0.409
food	drink	180	9	0.050	0.0

FIGURE K.10 – Table extraite de (Vulić *et al.*, 2017) avec des exemples d’association. Légende : Exemple (cue, response) pairs of free word association from the USF data set. #G stands for the number of participants serving in the group norming the word, while #P denotes the number participants producing a particular response. FSG= Forward association strength, BSG Backward