Figurative language processing: A developmental and NLP Perspective

Mila Vulchanova

Norwegian University of Science & Technology

mila.vulchanova@ntnu.no

Valentin Vulchanov

Norwegian University of Science & Technology

valentin.vulchanov@ntnu.no

Abstract

It is now common to employ evidence from human behaviour (e.g., child development) for the creation of computational models of this behaviour with a variety of applications (e.g., in developmental robotics). In this paper we address research in the comprehension and processing of figurative (non-literal) language in highly verbal individuals with autism in comparison with age- and language level-matched neuro-typical individuals and discuss critically what factors might account for the observed problems. Based on this evidence we try to outline the strategies used by human language users in understanding non-literal/non-compositional expressions and proceed to identifying possible solutions for automated language systems in the domain of idiomatic expressions.

1. Introduction

Figurative or non-literal language is a pervasive phenomenon in every-day human communication. It covers a wide range of expressions, such as idioms, metaphors, irony and jokes, hyperbole, indirect requests, as well as other stereotyped expressions, such as cliches. A recent study investigating the incidence of non-literal expressions in e-mails written by young people found that 94,30% of the e-mails included at least one non-literal statement and participants used on average 2,90 non-literal expressions per e-mail (Whalen, Pexman & Gill, 2009). Unlike literal language, where the interpretation depends on computing the meaning of each of the expression constituents, figurative expressions seem to require additional operations in order to arrive at the intended meaning. Furthermore, the competences and skills associated with figurative language mastery take much longer to develop than word knowledge (vocabulary) or core grammar. In typical language development, children appear to demonstrate appreciation for figurative expressions, such as idioms, at some point in the school years (Nippold, 1998; 2006; Nippold & Duthie, 2003; Cain et al., 2009, Levorato & Cacciari, 1995). Interestingly, this ability patterns in a way similar to the emergence of dimensionality in language competences and skills, as established recently in a large-scale cohort study covering pre-school to early school ages (LARRC, 2015). In that study, three clear dimensions of language competence (vocabulary, grammar and discourse) are first distinguished around third grade at school. It is still unclear, however, whether the developmental trajectory displays a linear trend over time (Nippold, 1998, 2006) or a quadratic trend peaking before adolescence, with less change afterwards (Kempler et al., 1999; Vulchanova et al., 2011; Laval & Bernicot, 2002). More intriguingly, figurative language skills, by taking longer to acquire, also manifest vulnerability both in developmental deficits and across the life-span. Research in typical ageing demonstrates that older adults produce fewer idioms and benefit more from cueing than younger speakers (Conner et al., 2011). Findings from acquired deficits, such as aphasia demonstrate impaired idiom comprehension (Milburn, Warren & Dickey, 2018). Problems with figurative language have been systematically documented in developmental deficits, such as autism spectrum disorder (ASD) (Volden & Phillips, 2010; Ramberg, Ehlers, Nydén, Johansson and Gillberg, 2011). Recent studies establish failure to understand pragmatic, non-literal aspects of language, such as metaphors, idioms and other forms of figurative language, even when structural language may appear to be intact (Gold and Faust, 2010; for a

Keywords: Figurative language, NLP, autism

comprehensive critical review of converging evidence from existing research see Vulchanova et al., 2015).

From a developmental and cognitive perspective the question then is what makes figurative (non-literal) language more challenging and open to such vulnerability. Given this vulnerability and complexity, a second related question is what can explain the high prevalence of non-literal language in discourse. Below we briefly present current accounts of figurative language processing and the main factors that impact on figurative language interpretation.

2. Accounts of figurative language processing

Unlike literal language, figurative language is non-transparent. It requires the hearer to go beyond the meaning of the individual constituent words in order to decode the speaker's intended meaning. Idioms, for instance, have lost their original semantic motivation, and need to be stored as multi-word expressions, very much like the way speakers store lexical items in long-term memory for the purposes of access and retrieval. Thus, expressions, such as *hit the sack* (literally «go to bed») or *kick the bucket* (literally «die») cannot be decoded solely based on co-composing the meaning of the verb (*hit, kick*) and its complement (*sack, bucket*), and the relationship between the head verb and its complement is not the same as between that same verb and an argument this verb subcategorises for (Nunberg et al., 1994). Yet, many idioms allow for partial analysability, in that one of its components functions according to its typical collocational environment, while it is the other constituent which requires a metaphorical interpretation, e.g., «the question» in *pop the question*. It has also been suggested that even in completely non-transparent expressions, such as *kick the bucket*, the head verb «kick» preserves its aspectual features preventing the acceptability of sentences, such as «??John lay kicking the bucket due to his chronic illness» or using an adverb which is inconsistent with the punctual feature of *kick* (Glucksberg, 1991; Hamblin & Gibbs, 1999). These properties of idioms have given rise to two types of accounts, non-compositional accounts, which acknowledge the need to store and retrieve idioms as multi-word chunks, and compositional accounts, which focus on the possibility of individual constituents to affect the interpretation or usage of the idiom.

2.1. Non-compositional accounts

A variety of accounts assume that idioms share similar properties, and are processed as, lexical items or multi-chunk words. Such approaches build on observations that the meaning of idioms is not a function of the meaning of their individual constituents and that language users need to acquire and store these expressions very much like words. In the linguistics tradition, this non decomposability issue has been addressed in numerous studies highlighting the impossibility to generate the expression following the rules of phrase structure and further to modify or change the structure (e.g., derive passives, insert modifiers etc. (Chomsky, 1980; Nunberg et al., 1994; Jackendoff, 2002). Non-compositional approaches differ on a range of parameters and specific assumptions. Thus, some approaches adopt single step processing, while others assume step-wise processing. According to the *standard pragmatic approach*, the first step in processing involves activating the literal meanings associated with the constituent words in the expression, only arriving at the intended figurative non-literal interpretation at the second step (Grice, 1975). In contrast, the *direct access model* assumes that there is no need for the initial activation of the literal meaning(s) (Gibbs, 1994). Instead, the figurative meaning is retrieved directly, following cues from the linguistic and other (e.g., communicative) context of the expression. The lexical nature of idioms has been recognised also in the *lexical representation hypothesis* (Swinney & Cutler, 1979), where idioms are assumed to be stored as lexical items, which can be retrieved fast, at the same time engaging in a second parallel process of lexical decomposition.

2.2. Compositional accounts

In contrast, compositional approaches rest on the assumption that idiom processing involves, at least at some level, de-composition. Thus, Hamblin & Gibbs (1999) insist on idiom decomposability, suggesting that idiom interpretation depends on identifying the individual constituents of the expression. Such approaches recognize that the processing and understanding of idioms cannot be reduced to lexical access or lexical retrieval only (Cacciari and Tabossi, 1988; Gibbs, 1992; Vega-

Moreno, 2001). A similar, albeit technically different, approach, is the *configuration hypothesis*. On this approach, idioms are represented in a distributed way and they are processed as complex expressions, very much like other instances of similar syntactic complexity (Cacciari and Tabossi, 1988).

2.3. The hybrid model

There is evidence in current research supporting both non-compositional and compositional approaches. The central question addressed in most studies is whether at all, in order to arrive at the target interpretation, speakers retrieve the literal aspects of constituents or by-pass this step, and retrieve the target non-literal interpretation. In a critical review of existing research and based on an experimental study, Titone & Connine (1999) propose a *hybrid model* for idiom processing where the main factor is idiom decomposability. Thus, idioms are assumed to function simultaneously as arbitrary associations between phonological form and meaning (like words), and compositional phrases. Processing of idiomatic expressions will then depend on the degree to which the idiom is inherently analyzable into constituent parts. It deserves mention here that compositionality of interpretation and the extent to which there is a need for literal processing have been brought to bear in models of the processing of other types of figurative language, for instance verbal irony and jokes (Dews & Winner, 1999).

3. Factors in figurative language processing

It becomes clear that the processing and interpretation of idioms depends on a number of key properties. Following Nunberg et al. (1994), three main factors have been identified, compositionality, conventionality and transparency. *Compositionality* applies to the degree to which the expression can be analysed into constituent parts or put differently, whether the expression reflects any syntactic and semantic structure. For instance, for an idiom comprising a head verb and its object (*spill the beans*), the question is whether this verb phrase structure is accessible to the speaker, and is it likely that speakers are using it for the processing of the idiom. *Conventionality* applies to the degree to which the expression has become lexically encoded and part of the lexical inventory of the language at hand. Conventionalisation is a natural process in language evolution and reflects the societal forces in language practice, whereby certain forms become adopted by the community of speakers by virtue of social agreement (de Saussure, 1964). *Transparency* applies to the extent to which the semantic motivation of the expression is evident. Even though these properties can be stated independently, they are not easily operationalisable, and quite often overlap. For instance, the difference between compositionality and transparency is often obscured by semantic judgements on whether the target interpretation is easily accessible or not. Furthermore, other criteria have been shown to play a role. For instance, *familiarity*, as established in population norming studies has a documented effect on processing latencies, as well as accuracy. Since language learning and use largely depend on frequency of exposure to linguistic input, frequency can be assumed to play a role in the processing of figurative expressions. Idiom familiarity and the frequency of constituent words interact, as evidenced in studies where the magnitude of the idiom familiarity effect seems to be diluted when the idiom contains low-frequency words as constituents (Cronk, Lima & Schweigert, 1993). Importantly, both the frequency of the individual constituents and their collocational frequency would play a role in how fast speakers access the target meaning, whether they retrieve literal meanings at all, and whether this process is marked by a competition between target figurative and literal interpretation. We address this question in more detail in 5. Below.

Conventionality/novelty often correlates highly with familiarity and frequency. However, there is no consensus on what test can be used as an objective measure of familiarity and how it can be operationalised (cf. Thibodeau, Sikos & Durgin, 2017 for a discussion). While some authors have used subjective measures, such as e.g., «the perceived experience with the metaphor» (Blasko & Connine, 1993), others suggest that frequency (measured in web corpora) can be used as an objective measure due to its high correlation with familiarity (Thibodeau & Durgin, 2011). In addition, conventionality and familiarity often do not yield clear independent effects in experimental research (Dulcinatti, Mazzarella, Pouscoulous & Rodd, 2014). Thus, a central methodological issue in research on figurative language is how to operationalize the factors that play a role in idiom processing and how to

establish objective measures to be used in experimental designs or modeling. Needless to say, a final crucial factor in idiom processing, is the presence of biasing *context*. Numerous studies have found effects of context which might bias for the target figurative interpretation of the expression or not.

4. Findings from research in ASD

The diversity and complexity of factors involved in the processing and comprehension of figurative language may be specifically challenging in developmental disorders, such as autism. Problems in this domain are well-attested (Tager-Flusberg, 2006; Volden & Phillips, 2010; Vulchanova, Saldaña, Chahboun & Vulchanov, 2015), however, their source remains largely controversial. In a series of specifically designed studies, we investigated performance on figurative language tasks (both idioms and metaphors) in highly verbal individuals with autism in comparison with IQ- and language ability-matched neuro-typical individuals. The participants in those studies came from two age groups, 10-12 (children) and young adults in the range 16 – 22 years analysed in a cross-sectional design. The two age ranges and the cross-sectional design were included specifically to establish possible developmental trajectories in both controls and experimental group.

Our main findings can be summed up in the following way. The main problems encountered by the participants with autism were primarily reflected in significantly greater reaction latencies in comparison to controls. The participants with autism performed at adequate levels of accuracy, though still displaying poorer responses in comparison to controls. Another major finding is the different developmental trajectories between the experimental groups and controls: young adult participants with autism performed at the level of control children, but better than children with autism, as evidenced by main effects of Age and Group in our results. We also have evidence of different underlying strategies in the processing of figurative language and in text comprehension.

A main finding in that research is that young adults with autism are less accurate than adults without autism. A valid question then is what types of errors are they making. The results in Chahboun et al. (2016b) suggest that the responses they provide are more literal. In this study a difference in degree of literalness was observed in response accuracy. The model revealed a main effect of group (control/ASD) ((χ^2 (1, 26) = 5.22, p = .022), with more literal responses by participants with autism, and a marginally significant difference in accuracy between Age (children/young adults) ((χ^2 (1, 26) = 3.51, p = .06). Furthermore, a two-way interaction between age and group was observed (χ^2 (1, 26) = 4.89, p =.02). Multiple comparisons with Tukey contrasts revealed that this interaction was due to a significant difference between control young adults and young adults with autism (p =.015), where the young adults with autism were converging on more literal responses than their typically developing peers. Thus, the younger participants and participants with autism in our study interpreted the stimuli more often literally than older participants and controls. These data provide support for other findings in research on young children and individuals with autism documenting a tendency for literal interpretation (Mitchell, Saltmarsh & Russell, 1997). Data from the same study further suggest that younger participants and participants with autism have specific problems with the idioms with greater decomposability, but no similar problems were observed with novel decomposable metaphors of literal expressions. This comes to suggest that idiom decomposability interferes in certain ways with processing and interpretation, unlike other decomposable expressions. These data find support in a recent ERP study of idiom comprehension in Chinese, whereby decomposable expressions, both idioms and free (literal) expressions elicited significantly greater ERP responses than the non-decomposable idioms.

5. Figurative language and NLP

On the backdrop of factors involved in the processing and comprehension of non-literal language in typical individuals, and the problems observed in highly verbal individuals with autism, a possible approach needs to look at what features of the expression might trigger literal (compositional) strategies, thus procrastinating the target figurative interpretation. We aim to outline under what conditions this is more likely to happen.

The central factor that needs to be considered is idiom decomposability and to what extent properties of the constituent words might trigger competition between a literal and target figurative

meaning. Some studies have used the notion of *semantic plausibility* suggesting that, in the absence of a biasing context, both a literal and a figurative interpretations are equally plausible. Thus, for instance the idiom *pull someone's leg* is equally plausible in a direct literal way, and figuratively. Other expressions do not easily yield such interpretations. For instance, *I am a bit under the weather today* cannot possibly make sense on a literal interpretation. It is to be expected that only the semantically plausible idioms may trigger competition between the literal and the figurative meaning, whereas the less plausible ones will directly cue the target figurative meaning, which is also the only plausible interpretation to be accessed. Another approach would be to assess the semantic similarity or closeness between the two available interpretations, the literal and the idiomatic (Milburn, 2017). This type of approach holds promise in circumventing issues arising from the need to categorise expressions according to decomposability, conventionality, transparency and other parameters. On this approach, frequency can be easily added in the equation to assess the collocational probability of one part of the expression co-occurring with the other part in comparison to collocational alternatives, e.g., the same word co-occurring with other lexical items.

To give a concrete example of an idiom like *kick the bucket*, one can estimate the probability of the noun phrase "the bucket" co-occurring with the head verb "kick" against the probability of the same verb co-occurring with other possible fillers of the complement position. The main measure can be **cloze probability** of dependent constituent, and we can assume that degree of activation of possible candidates for the verb complement position will depend on the ratio of cloze probabilities. In the case of idioms, this may be the ratio between the NP filler in the idiom and the most frequent literal filler of the argument position expressed in formula like ClProbMax **NP(VP)** / CP **NP(idiom)**. For instance, counts of these in the case of *kick the bucket* reveal that the most frequently occurring filler in the context of kick is "the ball" with native speaker cloze probability counts at 54%, and estimates from on-line corpora at 14.58%. Concerning the collocational frequency of "the bucket" as argument filler in the idiom, counts vary depending on corpus. A search in CoCA gives a value of 5.7 for **ball/bucket** suggesting that ball is by far more frequent after *kick*. Since corpora produce different results, cloze probabilities can be estimated in norming studies with native speakers, especially given the high correlation between cloze probabilities measured in sentence completion tasks with native speakers and in on-line (web) corpora (Hammerås, 2017).

We can then assume that the likelihood that literal interpretations may be activated can be measured as the value of the ratio between the cloze probabilities of the two argument filler candidates, and bigger values (according to the formula above) will lead to greater competition between the literal and figurative meaning as a result of more likely literal activation of head word/verb. This type of approach can be tested experimentally in a controlled design with carefully selected stimuli.

6. Conclusions

In this paper we have addressed issues arising from the factors which impact on the processing of figurative language against common assumptions and accounts. Based on evidence from behavioural research with idiom and metaphor processing in autism and typical individuals, we have proposed an approach which captures the common problems encountered by special populations, and often children, in the processing of non-literal language, at same time offering a solution to how to operationalize idiom processing in a measurable and meaningful way, consistent with how the human brain may be handling the task. Such an approach is testable, and as such, can be useful for computational modeling of natural language processes.

References

Blasko, D. G., & Connine, C. M. (1993). Effects of familiarity and aptness on metaphor processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19(2),* 295-308.

Cacciari, C. & Tabossi, P. (1988). The comprehension of idioms. *J Mem Lang*., 27:668–83

Cain, K., Towse, A. S., and Knight, R. S. (2009). The development of idiom comprehension: an investigation of semantic and contextual processing skills. *J. Exp.Child Psychol. 102*, 280–298. doi:10.1016/j.jecp.2008.08.001

Chahboun, S., Vulchanov, V., Saldaña, D., Eshuis, H. & Vulchanova, M. (2016b). Can You Play with Fire and Not Hurt Yourself? A Comparative Study in Figurative Language Comprehension between Individuals with and without Autism Spectrum Disorder. *PLoS ONE* 11(12): e0168571. https://doi.org/10.1371/journal.pone.0168571.

Chomsky, N. (1980). *Rules and representations*. New York: Columbia University Press.

Conner, P. S., Hyun, J., O'Connor Wells, B., Anema, I., Goral, M., Monéreau-Merry, M., Rubino, D., Kuckuk, R. & Obler, L. K. (2011). Age-related differences in idiom production in adulthood. *Clinical Linguistics & Phonetics, 25(10)*, 899-912.

Cronk, B., Lima, S. & Schweigert, W. (1993). Idioms in sentences: Effects of frequency, literalness, and familiarity. *Journal of Psycholinguistic Research, 22*, 59. https://doi.org/10.1007/BF01068157

de Saussure, F. (1964). *Course in General Linguistics*, 2nd impression. London: Peter Owen.

Dews, S. & Winner, E. (1999). Obligatory processing of literal and nonliteral meanings in verbal irony. *Journal of Pragmatics 31*, 1579-1599

Dulcinati, G., Mazzarella, D., Pouscoulous, N. & Rodd, J. (2014). Processing metaphor: The role of conventionality, familiarity and dominance. *UCL Working Papers in Linguistics, 26*, 72-88.

Gibbs, R. W. (1992). Categorization and metaphor understanding. *Psychol. Rev. 99*, 572–577. doi: 10.1037//0033-295x.99.3.572

Gibbs R. (1994). *The poetics of mind: Figurative thought, language and understanding.* Cambridge: Cambridge University Press.

Glucksberg, S. (1991). Beyond literal meanings: The psychology of allusion. *Psychological Science 2*: 146-152.

Gold, R. & Faust, M. (2010). Right hemisphere dysfunction and metaphor comprehension in young adults with Asperger syndrome. *J. Autism Dev. Disord. 40*, 800–811. doi: 10.1007/s10803-009-0930-1

Grice P. (1975) Logic and conversation. In Cole P, Morgan J (eds.). *Syntax and Semantics 3: Speech Acts*. New York: Academic Pressp, 41-58.

Hamblin, J. & Gibbs, R. (1999). Why You Can't Kick the Bucket as You Slowly Die : Verbs in Idiom Comprehension. *Journal of Psycholinguistic research*, 28(1): 25-39.

Hammerås, M. L. (2017). Probing sensitivity to argument structure in two proficiency level groups - an exploratory study with Norwegian learners of English. MA thesis, NTNU. (https://www.ntnu.edu/documents/1535402/35615794/Hammer%C3%A5s2017/0c92037d-7784-4bf5-9d76-53bb1872e861)

Jackendoff, R. (2002). What's in the lexicon? In: Nooteboom S, Weerman F, Wijnen F (eds) *Storage and computation in the language faculty*. Dordrecht: Kluwer, 23–58.

Kempler,D.,VanLancker,D.,Marchman,V.,andBates,E.(1999).Idiom Comprehension in children and adults with unilateral damage. *Dev. Neuropsychol*. 15, 327-349. doi:10.1080/87565649909540753

LARRC. (2015), The dimensionality of language ability in young children. Child Development, 86: 1948–1965. doi: 10.1111/cdev.12450

Laval, V., and Bernicot, J. (2002). Tu es dans la lune: understanding idioms in French speaking children and adults. *Pragmatics 12*, 399–413.

Levorato, M. C. & Cacciari, C. (1995). The effects of different tasks on the comprehension and production of idioms in children. *J. Exp. Child Psychol. 60*, 261–283. doi: 10.1006/jecp.1995.1041

Milburn, E. (2018). *The effects of meaning dominance and meaning relatedness on ambiguity resolution: Idioms and ambiguous words*. Doctoral Dissertation, University of Pittsburgh.

Milburn, E., Warren, T. & Dickey, M. W. (2018). Idiom comprehension in aphasia: Literal interference and abstract representation. *Journal of Neurolinguistics* (https://doi.org/10.1016/j.jneuroling.2018.02.002)

Mitchell, P., Saltmarsh, R. & Russell, H. (1997). Overly Literal Interpretations of Speech in Autism : Understanding That Messages Arise from Minds. *The Journal of Child Psychology and Psychiatry, 38(6)*, 685–91.

Nippold, M. (1998). *Later Language Development: The School-Age and Adolescent Years.* 2nd Edn. Austin, TX: Pro-Ed.

Nippold, M. A. (2006). "Language development in school-age children, adolescents and adults," in *Encyclopedia of Language and Linguistics* (Vol.6), 2nd Edn. Ed K. Brown (Oxford, UK: Elsevier Publishing), 368–372.

Nippold, M. A., and Duthie, J. K. (2003). Mental imagery and idiom comprehension: a comparison of school-age children and adults. *J. Speech Lang. Hear. Res. 46*, 788–799. doi:10.1044/1092-4388(2003/062)

Nunherg, G., Sag, 1., & Wasow, T. (1994.) "Idioms". *Language, 70, 3*, 491-538.

Ramberg, C., Ehlers, S., Nydén, A., Johansson, M. & Gillberg, C. (2011). Language and pragmatic functions in school-age children on the autism spectrum. *International Journal of Language & Communication Disorders, 31*, 387-413.

Swinney D. & Cutler A. (1979). The Access and Processing of Idiomatic Expressions. *Journal of Verbal Learning and Verbal Behavior, 18*, 523–34.

Tager-Flusberg, H. (2006). Defining language phenotypes in autism. *Clinical Neuroscience Research, 6*, 219-224.

Thibodeau, P. H., & Durgin, F. H. (2011). Metaphor aptness and conventionality: A processing fluency account. *Metaphor and Symbol, 26(3)*, 206-226.

Thibodeau, P. H., Sikos, L., & Durgin, F. H. (2017). Are subjective ratings of metaphors a red herring? The big two dimensions of metaphoric sentences. *Behavior Research Methods*, 1-14.

Titone D. & Connine C. (1999). On the compositional and noncompositional nature of idiomatic expressions. *J Pragmat, 31(12)*, 1655–74.

Vega-Moreno, R. E. (2001). Representing and processing idioms. *UCLWPL 13*, 70–109.

Volden, J. & Phillips, L. (2010). Measuring Pragmatic Language in Speakers with Autism Spectrum Disorders: Comparing the Children's Communication Checklist 2 and the Test of Pragmatic Language. *American Journal of Speech-language Pathology, 19 (3)*, 204-212.

Vulchanova, M., Vulchanov, V. & Stankova, M. (2011). Idiom comprehension in the first language : a developmental. *Vigo International Journal of Applied Linguistics, 8*, 207–34.

Vulchanova M., Saldaña D., Chahboun S. & Vulchanov V. (2015). Figurative language processing in atypical populations : the ASD perspective. *Frontiers in Human Neuroscience, 9(24)*, 1–11.

Whalen, J. M., Pexman, P. M & Gill, A. J. (2009) “Should Be Fun—Not!” Incidence and Marking of Nonliteral Language in E-Mail. *Journal of Language and Social Psychology, 28, 3*, 263-280.