Beyond the Senses: How Self-Directed Speech and Word Meaning Structure Impact Executive Functioning and Theory of Mind in Individuals With Hearing and Language Problems

Thomas F. Camminga¹,², Daan Hermans¹,², Eliane Segers¹ and Constance T. W. M. Vissers¹,²

¹ Behavioural Science Institute, Radboud University, Nijmegen, Netherlands, ² Royal Dutch Kentalis, Sint-Michielsgestel, Netherlands

Many individuals with developmental language disorder (DLD) and individuals who are deaf or hard of hearing (D/HH) have social–emotional problems, such as social difficulties, and show signs of aggression, depression, and anxiety. These problems can be partly associated with their executive functions (EFs) and theory of mind (ToM). The difficulties of both groups in EF and ToM may in turn be related to self-directed speech (i.e., overt or covert speech that is directed at the self). Self-directed speech is thought to allow for the construction of non-sensory representations (i.e., representations that do not coincide with direct observation). Such non-sensory representations allow individuals to overcome the limits set upon them by the senses. This ability is constrained by the development of word meaning structure (i.e., the way words are understood). We argue that the greater ability to construct non-sensory representations may result in more enhanced forms of EF and ToM. We conclude that difficulties in EF, ToM, and social–emotional functioning in those with hearing and language problems may be accounted for in terms of word meaning impairments. We propose that word meaning structure and self-directed speech should be considered in assigning EF and ToM treatments to individuals with DLD and those who are D/HH.

Keywords: self-directed speech, inner speech, word meaning structure, language, executive functioning, theory of mind, developmental language disorder, deaf and hard of hearing

INTRODUCTION

Individuals with developmental language disorder (DLD) and those who are deaf or hard of hearing (D/HH) often have social–emotional problems in addition to their learning problem (Dammeyer, 2010; Marschark et al., 2015; Hubert-Dibon et al., 2016; Conti-Ramsden et al., 2018). Both groups have lower self-rated self-esteem, higher risk of being bullied, and more problems with maintaining peer relations (e.g., Fellinger et al., 2015 and van den Bedem et al., 2018, 2019). In addition, they show increased signs of aggression, depression, and anxiety (e.g., Kvam et al., 2006; Theunissen et al., 2014, and Brownlie et al., 2016).
In both groups, these social–emotional problems have been linked to difficulties in executive functioning (EF; e.g., Pauls and Archibald, 2016 and Botting et al., 2017) and theory of mind (ToM; e.g., Meristo et al., 2007 and Nilsson and de López, 2016). Recently, the claim has been made that these EF and ToM difficulties may be due to delays in the development of self-directed speech (Aziz, 2015; Vissers and Hermans, 2018; Mulvihill et al., 2019; Vissers et al., 2020). Here, we extend this hypothesis, by arguing that language problems are reflected by a limited understanding of word meanings, constraining the potential of self-directed speech in supporting EF and ToM.

THE INTERPLAY BETWEEN SELF-DIRECTED SPEECH, EF, AND TOM IN INDIVIDUALS WITH HEARING AND LANGUAGE PROBLEMS

DLD is attributed to individuals who are delayed in their language development in the absence of a known biomedical etiology (Bishop et al., 2017). DLD is a heterogeneous disorder that may be characterized by various underlying neuropsychological deficits (Tomas and Vissers, 2019). It has a prevalence of 7–14% in children younger than 5 years (Law et al., 2017). Many individuals who are D/HH are impaired in all aspects of language relative to their normal peers and even in some cases to children with DLD (Tomblin et al., 2015; de Hoog, 2017). In addition to their language problems, both groups have difficulties in EF and ToM (e.g., Hintermair, 2013; Vugs et al., 2014). For example, EF problems are three to five times more common in children that are D/HH compared to typically developing children (Hintermair, 2013). Interestingly, EF and ToM problems are generally restricted to D/HH children with hearing parents rather than those with D/HH parents (Schick et al., 2007; Hall et al., 2017), suggesting that their language problems result from a mismatch between their perceptual abilities and those of their family (Hall et al., 2019). Thus, sharing language (or communication more generally), be it spoken or signed, appears to be an important factor in the development of EF and ToM. This is corroborated by longitudinal relationships between language, on the one hand, and EF and ToM, on the other (Milligan et al., 2007; Slot and von Suchodoletz, 2018). Note that these relations are likely bidirectional, as EF and ToM have also been shown to support language development (e.g., Loosli et al., 2012; for children with DLD see Sikora et al., 2019). Here, we theoretically explore only the mechanisms underlying the first direction of causality.

The dominant view of EFs holds that they are “general-purpose control mechanisms that modulate the operation of various cognitive subprocesses and thereby regulate the dynamics of human cognition (Miyake et al., 2000, p. 50).” In social–emotional functioning, EFs help individuals, for example, to restrain impulsive or inappropriate actions, to shift their attention away from negative stimuli, and to modify their goals and plans in the light of the needs, goals, impulses, and emotions of others (Vissers and Hermans, 2018). Three EFs, with distinct neuroanatomical substrates, are generally considered to be the core EFs: working memory updating (updating), inhibition of prepotent responses (inhibition), and mental set shifting (shifting).

ToM is defined as the ability to understand the behavior of others in mental terms (Premack and Woodruff, 1978). It is preceded by a complex developmental path that includes several precursors, such as the capacity for imitation, joint attention, and emotion recognition and understanding. Four-year-olds may learn that individuals can have false beliefs, which is taken as a hallmark of ToM (Wellman et al., 2001). Starting from the age of 7 years, children may learn to distinguish what is said from what is meant (e.g., sarcasm). ToM supports prosocial behavior, by importing considerations of the thoughts and feelings of other people into the decision-making process. Two dimensions of ToM with distinctive neuroanatomical underpinnings can be discerned (Westby and Robinson, 2014). Cognitive ToM refers to reflections based on thoughts, beliefs, and intentions, whereas affective ToM concerns reflections on feelings and emotions (e.g., Dvash and Shamay-Tsoory, 2014). These reflections may be directed to one’s own mental states (intrapersonal ToM) or those of others (interpersonal ToM; Tine and Lucariello, 2012).

A potential explanatory account for the difficulties of individuals with language problems in EF and ToM was suggested by Vygotsky and Luria (1930/1994). These authors traced the origins of self-directed speech to the social dialogue. They observed in their experiments that when children tried to get a desired object that was out of reach, they asked the experimenter for help. When the experimenter left the room, however, the children continued speaking about the object and their own behavior toward it, but now to themselves. In children around the age of 6 years, self-directed speech typically starts to internalize (it “goes underground;” Vygotsky, 1934/1986, p. 34) until it finally becomes silent (i.e., inner speech; Vygotsky, 1934/1986; Bivens and Berk, 1990; Damianova et al., 2012). In children with language problems, this internalization process appears to be delayed (i.e., inner speech and private speech emerge at a later age; Lidstone et al., 2012; Aziz et al., 2017), and they draw upon it to a lesser degree in planning (Kuvalja et al., 2014; Larson et al., 2019).

Self-directed speech—including its equivalent in sign language, self-directed signing—is universal among humans (e.g., Al Namlah et al., 2006; Zimmermann and Brugger, 2013; Thibodeaux et al., 2019), although its frequency and manner of application may vary between individuals and tasks (Alderson-Day and Fernyhough, 2015). The development of self-directed speech is thought to be completely intertwined with that of other cognitive functions such as EF and ToM (e.g., Newton and de Villiers, 2007; Lidstone et al., 2010, 2012). More precisely, self-directed speech can, under the influence of the social environment, be synthesized with (precursors of) EF and ToM into functional systems, which have properties that none of these cognitive functions have on their own (Vygotsky and Luria, 1930/1994; Fernyhough, 2010; Toomela, 2016). For example, false belief understanding has been hypothesized to emerge as a result of (social) activity-driven developments (e.g., incidental learning; Marschark and Knoors, 2014) in cognitive functions as diverse as elementary forms of ToM (e.g., joint attention;
Tomasello, 2019), EF (e.g., shifting between perspectives of self and other), and language (as a representational format; Frye et al., 1995; Fernyhough, 2008).

**BEYOND THE SENSES: NON-SENSORY REPRESENTATIONS AND WORD MEANING STRUCTURE**

According to Vygotsky and Luria (1930/1994) and Toomela (2016), the development of self-directed speech grounds new ways of relating to the mind and the external world. Words—or more generally, symbols (including signs)—are linked to referents, which are mental images that correspond (directly or indirectly) to an aspect of the world (de Saussure, 1966). Importantly, symbols can be brought in contexts that are not possible for their referents (Toomela, 2016). However, even though symbols can be used in the absence of their referents, symbol and referent still form a holistic unity in the mind—the activation of either the symbol or its referent results in the activation of the other (de Saussure, 1966). Consider, for example, the sentence “a giraffe is having a picnic at the bottom of the ocean.” This fictional sentence may evoke a mental image of a giraffe even if there is no giraffe in the direct environment. Thus, humans are able to temporarily ignore the immediate sensory present (Vygotsky and Luria, 1930/1994), and to construct non-sensory representations: mental descriptions or images of objects or events that do not coincide with, or even contradict, the immediately sensed present. Therefore, self-directed speech allows to represent the non-sensory world (i.e., aspects of the world that cannot be observed through the sensory organs; Toomela, 2016), as well as events and phenomena that are observable in principle, but not at the moment.

These properties of self-directed speech have important consequences for EF and ToM. However, the potential for constructing non-sensory representations is constrained by language development. In addition, based on earlier suggestions by Vygotsky (1934/1986) and Luria (1976), Toomela (2003, 2020b) proposed that word meaning structure (i.e., the way words are understood) may be especially relevant in this respect. Word meaning structure is explicitly related to qualitative developments in the potential for articulating certain types of verbal contents and fits with Vygotsky’s functional systems approach. Thus, this article focuses on word meaning structure, without denying the roles of other language aspects (e.g., vocabulary, syntax, and pragmatics) in EF and ToM that have been pointed out by other researchers (e.g., Harris, 2005; Milligan et al., 2007; Müller et al., 2009).1

Toomela (2003, 2020b) distinguishes five word meaning stages (Table 1). They are organized hierarchically, meaning that later stages emerge on the basis of earlier stages. Moreover, word meaning development is domain-specific, meaning that developments in one area do not guarantee improvements in other domains. In the first word meaning stage, that of *syncretic concepts*, words have no fixed relation to their referent. A child may use a single word to refer to different aspects of a situation (an object, a property, or the whole situation), depending on the context. Next, in the stage of *object concepts*, two classes of words are differentiated, namely, words that refer to objects and object-specific properties. In the third stage, *everyday concepts*, children can implicitly learn all grammatical classes, allowing children to describe situations (i.e., relations between objects). In this stage, all aspects of the sensory world can be represented, as well as aspects of the non-sensory world, and even references to the past and the future. Categories that are signified by everyday concepts still have fuzzy boundaries, meaning that things can belong to categories in different degrees. Categorizations at this stage are based mainly on perceptual similarity and everyday activities. Logical concepts, in contrast, are related to each other in a hierarchical taxonomy. Because of this hierarchical structure, logical concepts allow for a conscious differentiation of thought processes from the objects of thought and to group phenomena based on non-sensory properties. Moreover, logical concepts are characterized by categories with sharp, verbally defined boundaries that allow for abstract reasoning. Finally, *systemic concepts* embed these sharp, verbally defined categories within a broader system, as it can be realized that one object can belong to multiple categories, depending on the context (Figure 1).

**A VYGOTSKIAN PERSPECTIVE ON EF DEVELOPMENT: LINKS WITH THE STAGES OF WORD MEANING STRUCTURE**

Vygotsky and Luria (1930/1994) showed that symbols can mediate the influences of the environment on individuals, thereby allowing individuals to regulate their cognitive processes and behavior. Here we argue that word meaning structure may play a unique role in EF, by constraining the potential for self-directed speech to construct non-sensory representations. In order to understand the problems of those with language and hearing problems in EF, we will propose advancements in the components of EF, namely, updating, inhibition, and shifting (Miyake et al., 2000), which result from each new word meaning stage.

**Syncretic and Object Concepts**

Syncretic concepts and object concepts allow individuals to verbally label a stimulus. Verbal labels can be decoupled from their referents, resulting in an enduring trace that can be represented in working memory even in the referent’s absence (Luria, 1962/1980; Al Namlah et al., 2006; Müller et al., 2009). Labeling thus allows individuals to bring established action programs in novel situations. Moreover, labels single out essential, and inhibit inessential aspects of the environment (Luria, 1961; Toomela, 2002; Müller et al., 2004).

---

1The concepts of self-directed speech and word meaning structure should be clearly distinguished, because two individuals in the same word meaning stage may differ in the way they construct non-sensory representations in particular situations in self-directed speech.

2These stages concern verbal concepts (not preverbal, sensory-based concepts). However, for Vygotsky, the development of linguistic concepts is completely intertwined with cognitive development more broadly.

3Logical concepts correspond to Vygotsky’s (1934/1986) “scientific concepts.”
labels have been shown to enhance shifting abilities (Jacques, 2001).

**Everyday Concepts**

The opportunities for regulating behavior increase drastically with the emergence of everyday concepts. Luria (1962/1980) states that inner speech “plays an active part in […] singling out the aim of the action and providing a general scheme for it (p. 292).” In other words, self-directed speech can dictate an action plan that can be maintained in working memory in the face of changing environments. The inhibition of goal-irrelevant behaviors and information can be based on these verbal plans (Luria, 1962/1980). Verbal plans may mediate cases where behavioral patterns come into conflict. Interestingly, it is not until the age of 4 years that children can shift successfully on the flexible item selection task (FIST; Jacques, 2001) and the dimensional change card sorting (DCCS) task (Zelazo et al., 1996)—two tasks that involve shifting between conflicting verbal response rules.

**Logical and Systemic Concepts**

By thinking in logical and systemic concepts, individuals can represent more coherent, accurate, and precise verbal plans that should facilitate more efficient inhibition and shifting. For example, a verbal rule such as “I will not look at my phone for the next hour” presupposes a sharp delineation of the word “not,” which is not supported by everyday conceptual thinking. Indeed, in a task that involves the memorization of two separate lists of words, Toomela et al. (2020) found that participants thinking predominantly in logical concepts were less susceptible to interference than those thinking in everyday concepts. Moreover, in a second task, the logical-conceptual thinkers had fewer difficulties in regulating their behavior in the face of potential interference.

**A Vygotskian Perspective on TOM Development: Links with the Stages of Word Meaning Structure**

Many abilities linked to ToM, such as imitation and joint attention, can be coordinated without self-directed speech (Tomasello, 2019). However, thoughts and feelings cannot be observed directly through the senses (Premack and Woodruff, 1978). Therefore, explicit ToM—the ability to attribute mental states to oneself and others—requires, from the Vygotskian point of view (Vygotsky and Luria, 1930/1994; Toomela, 2016), the involvement of self-directed speech. Consequently, we argue that...
FIGURE 1 | A Vygotskian account of the social–emotional problems of individuals with hearing and language problems. Social emotional functioning can be explained in terms of executive functioning and theory of mind, which in turn are affected by non-sensory representations created in self-directed speech. Non-sensory representations are images or descriptions corresponding to external events that do not coincide with direct observation. The ability to construct non-sensory representations is constrained by the level of word meaning structure.

the development of the dimensions of ToM (cognitive, affective, interpersonal, intrapersonal) is constrained by the level of word meaning structure.

Syncretic and Object Concepts

Syncretic concepts allow children to label their sensory experiences (object concepts add precision). Regarding affective ToM, these may include bodily sensations associated with one's own emotions, and facial expressions and body posture in others. Overtly or covertly saying labels brings the referents into awareness (Kolk, 2012; Toomela, 2016). However, these labels can initially only be connected to referents associated with observable phenomena (Toomela, 2020a). Therefore, the meanings of early mental state words may differ from those used by older children and adults (Booth et al., 1997). Still, labeling may facilitate aspects of ToM that emerge before the age of 3 years (Westby and Robinson, 2014), such emotion recognition, altruistic behavior, and prediction of the behavior of others, by making important aspects in the body and the environment more salient.

Everyday Concepts

Fernyhough (2008) suggested that young children come to understand mental states, not through metacognitive inference, but by representing perceptual, epistemic, and affective perspectives of oneself and others in self-directed speech. The ability to verbally describe a situation from another person's perspective and to differentiate it from one's own can emerge on the basis of everyday concepts. For example, false belief understanding, a hallmark of cognitive ToM, emerges after the age of 4 years (Wellman et al., 2001). Linguistic devices such as complementation syntax (de Villiers and de Villiers, 2014) and contrastives (Wellman, 2014) may allow individuals to mentally differentiate between perspectives of oneself and others.

Logical and Systemic Concepts

Logical concepts, through their hierarchical organization, support individuals in distinguishing their thought processes from the objects of thought, which indeed appears to be difficult for young children (e.g., Lagattuta et al., 2015 and Flavell et al., 1986). Thus, logical concepts may allow individuals to distinguish logical and probable from illogical and improbable inferences of mental states and processes. This may facilitate aspects of higher-order ToM, such as comprehension of lies, sarcasm, and figurative language, which emerge from 7 years old (Westby and Robinson, 2014). By increasing awareness of mental processes, logical concepts allow individuals to influence mental states (both cognitive and affective) more effectively in themselves and others (Vygotsky, 1934/1986). For example, awareness of negative cognitive distortions, as is facilitated by cognitive therapy, allows individuals to mitigate their influences on well-being (Leahy, 2017).

EF AND TOM PROBLEMS IN RELATION TO WORD MEANING STRUCTURE IN INDIVIDUALS WITH HEARING AND LANGUAGE PROBLEMS

As word meaning belongs to the language system (Vygotsky, 1934/1986; Luria, 1962/1980), it seems likely that individuals with hearing and language problems are delayed in their word meaning development. If so, this may explain their problems in their EF and ToM.

Problems in EF

Deficits in EF are already present in preschoolers with DLD (Vissers et al., 2015) and preschoolers who are D/HH (Beer et al., 2014). Early EF deficits may be related to problems in regulating mental processes with verbal labels. Three-year-olds typically start subjecting their behavior to verbal plans. From
age, differences emerge between children with DLD and their typically developing peers on tasks that involve shifting between conflicting verbal response rules, such as the FIST (Roello et al., 2015) and the DCCS (Farrant et al., 2012). In older children and adolescents with hearing and language problems, EF deficits may be related to logical concept acquisition. Indeed, children with language problems have difficulties in operating with taxonomic information, a hallmark of logical conceptual thought (e.g., Marinellie and Johnson, 2002; Marschark et al., 2004, and Dosi and Gavriilidou, 2020).

Problems in ToM

Deficits in cognitive and affective forms of social understanding are already evident in preschoolers with DLD (Vissers and Koolen, 2016) and preschoolers who are D/HH (e.g., Meristo et al., 2007 and Wiefferink et al., 2013). From the age of 3 years, clear differences emerge between children with hearing and language problems and their peers on false belief tasks (e.g., Meristo et al., 2007 and Nilsson and de López, 2016), and sociodramatic play (e.g., Cornelius and American, 1990 and Brown et al., 1997). This suggests that they find it harder to see the world from another person’s perspective, indicating delayed everyday conceptual ToM development. No research exists yet on logical conceptual ToM in individuals with DLD and in those who are D/HH, but given the indications for problems with logical conceptual thinking, high performance in this domain seems unlikely.

CONCLUSION

To conclude, we have aimed at achieving a better understanding of the social–emotional difficulties of individuals with DLD and those who are D/HH. We have argued that these individuals may have a less advanced word meaning structure, resulting in a limited potential for self-directed speech to support EF and ToM with non-sensory representations.

Strong conclusions are prohibited by a paucity of direct research on the topic. Future experimental and longitudinal research should assess whether this mechanism may indeed account for the social–emotional and academic problems of individuals with language problems (and vice versa). Thereby, it will be important to assess whether problems occur at the level of word meaning development or at the level of applying the learned linguistic concepts in concrete tasks (i.e., self-directed speech). One clinically relevant way to test this mechanism is by assessing whether word meaning structure relates to the effectiveness of treatments for individuals with DLD or individuals who are D/HH. For example, young children may benefit most from verbal labeling training (Jacques, 2001), whereas older children may benefit more from verbal planning for EF development (Abdul Aziz et al., 2016) and sociodramatic play (Qu et al., 2015) and metacognitive training for their ToM development. Through its mechanistic explanation of the interplay between language, EF, and ToM, the present framework could be used to improve existing EF and ToM treatments for individuals with hearing and language problems and to assign these treatments to specific stages of (word meaning) development.

AUTHOR CONTRIBUTIONS

Based on earlier work by CV and DH, TC has created the theoretical model described in the text. All authors have contributed to finalizing the perspective article.

FUNDING

This work was supported by The Royal Dutch Kentalis.

ACKNOWLEDGMENTS

The authors would like to thank Aaro Toomela, Lidy Smit, and Stefan Gaillard for their comments on drafts of the manuscript.

REFERENCES

Abdul Aziz, S., Fletcher, J., and Bayliss, D. M. (2016). The effectiveness of self-regulatory speech training for planning and problem solving in children with specific language impairment. J. Abnorm. Soc. Psychol. 44, 1045–1059. doi: 10.1007/s10802-015-0115-7

Al Namlah, A. S. A.-N., Fernyhough, C., and Meins, E. (2006). Sociocultural influences on the development of verbal mediation: private speech and phonological recoding in Saudi Arabian and British samples. Dev. Psychol. 1, 117–131. doi: 10.1037/0012-1649.42.1.117

Alderson-Day, B., and Fernyhough, C. (2015). Inner speech: development, cognitive functions, phenomenology, and neurobiology. Psychol. Bul. 141, 931–965. doi: 10.1037/bul0000021

Aziz, S. A. (2015). Self-talk during planning and problem solving in young children with specific language impairment (Doctoral Thesis). University of Western Australia, Perth, WA, Australia.

Aziz, S. A., Fletcher, J., and Bayliss, D. M. (2017). Self-regulatory speech during planning and problem-solving in children with SLI and their typically developing peers. Int. J. Lang. Commun. Disord. 52:311–322. doi: 10.1111/1460-6984.12273

Beer, J., Kronenberger, W. G., Castellanos, I., Colson, B. G., Henning, S. C., and Pisoni, D. B. (2014). Executive functioning skills in preschool-age children with cochlear implants. ASHA 57, 1521–1534. doi: 10.1044/2014_JSLHR-H-13-0054

Bishop, D. V. M., Snowling, M. J., Thompson, P. A., Greenhalgh, T., and The CATALISE-2 consortium (2017). Phase 2 of CATALISE: a multinational and multidisciplinary Delphi consensus study of problems with language development: terminology. J. Child Psychol. Psychiatry Allied Discip. 58, 1068–1080. doi: 10.1111/jcpp.12721

Bivens, J. E., and Berk, L. E. (1990). A longitudinal study of the development of elementary school children’s private speech. Merrill-Palmer Q. 66, 443–463.
