

Cool Vendors in Speech and Natural Language

Published: 17 October 2018 **ID:** G00369018

Analyst(s): Anthony Mullen, Alexander Linden, Jessica Ekholm, Adrian Lee, Erick Brethenoux

Technology product managers must exploit recent advances in speech and natural language as they continue to reshape interactions and intelligence across interface design, content development and analytics.

Key Findings

- As speech-to-text and text-to-speech technologies gain prominence in user interfaces, approaches to security and identity become essential and a key point of differentiation.
- Deep learning continues to propel many speech and natural language capabilities forward — even short-form natural language generation.
- Graph technologies and visualizations become critical to present complex relationships between natural language concepts.
- Smaller vendors often underplay the role of NLP APIs in their first rounds of product development.

Recommendations

Technology product managers looking to maximize personal technology platforms to create future experiences should:

- Prioritize security of voice interactions given consumer concern over misuse and leverage metadata from voice to improve both service and security.
- Re-examine where new approaches with deep learning can improve workflow, from text mining to short-form content creation.
- Ensure at least a third of development effort is spent on graphical presentations, such as graphs, designed to reduce the complexity of dealing with and understanding the complexity of models and results.
- Ensure platforms have flexible APIs to stimulate wider adoption and foster innovation by third parties.

Table of Contents

Strategic Planning Assumptions..... 2

Analysis..... 2

 What You Need to Know..... 2

 DeepL (Formerly Linguee)..... 4

 Phrasee..... 4

 Pindrop..... 6

 SavantX..... 7

 Where Are They Now?..... 8

Gartner Recommended Reading..... 8

List of Figures

Figure 1. Enterprises Are Exploring a Wide Range of Speech and Natural Language Technologies..... 3

Strategic Planning Assumptions

By 2022, 50% of all short-form copy in marketing will be generated using natural language generation.

By 2023, at least one text-to-speech provider cloning voices will be fined €1 million for misconduct.

By 2021, the top three smart speaker companies will have biometric authentication and multiparty recognition.

By 2022, over 70% of all finance and insurance organizations will use speech analysis to reduce risk.

Analysis

This research does not constitute an exhaustive list of vendors in any given technology area, but rather is designed to highlight interesting, new and innovative vendors, products and services. Gartner disclaims all warranties, express or implied, with respect to this research, including any warranties of merchantability or fitness for a particular purpose.

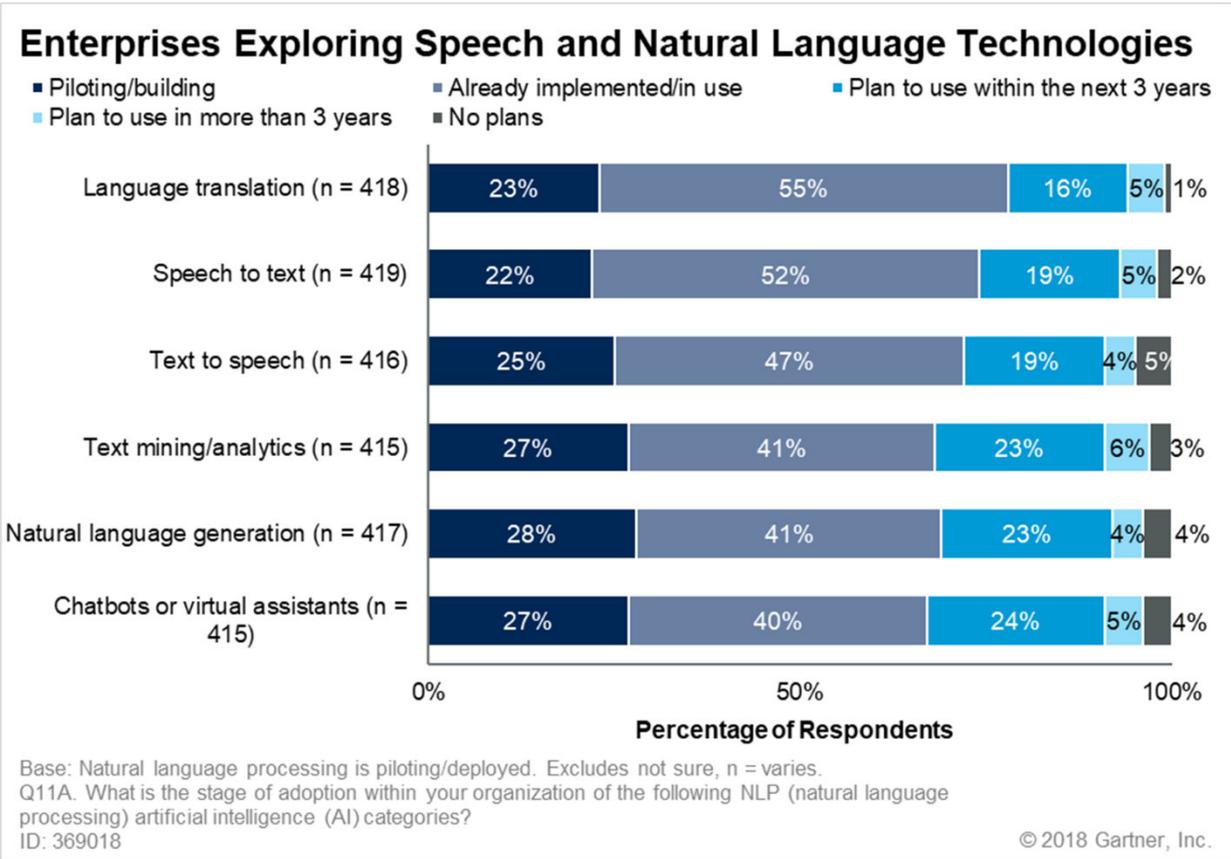
What You Need to Know

Where once organizations were content with rudimentary mining of text and sentiment analysis, we now see a much more diverse application of speech and natural language technologies across a wider variety of use cases. The tech titans and their consumer offerings, such as smart speakers

and virtual assistants, have, in part, helped stimulate technical advances, adoption and innovation by other players in a market that is now growing. These new technologies are helping reshape interactions between people and machines and represent a profound change — from people learning how to communicate with systems to systems learning how to communicate with people.

While chatbots are still a major focus for many, we found that, when we surveyed organizations in China, India, U.S., U.K. and Canada undertaking natural language processing pilots and deployments, they are investing across a wide range of activities. (See Figure 1.)

Figure 1. Enterprises Are Exploring a Wide Range of Speech and Natural Language Technologies



Source: Gartner (October 2018)

These new approaches in speech and natural language draw on a variety of techniques from semantic models, knowledge graphs, machine learning and deep neural networks. In particular, we see deep neural networks obtaining major improvements across translation, conversational systems, the interpretation and generation of speech and even short-form natural language generation.

Note: The Cool Vendors below are listed in alphabetical order.

DeepL (Formerly Linguee)

Cologne, Germany (www.deepl.com/en/home)

Analysis by Alexander Linden and Bern Elliot

Why Cool: DeepL announced in August 2017 the launch of DeepL Translator, a new translation service that has been cited¹ as superior to other machine translation offerings by Google, SYSTRAN, Baidu and others. DeepL Translator is offered free of charge. The tool can simply be used via an internet browser and has enjoyed rapid popularity in the past 12 months.

The company also delivered a more professional, commercially licensed, tool — DeepL Pro — in March 2018, which can handle bulk translations via a REST-API and allows for integration into computer-aided translation (CAT) tools.

Like many others, DeepL utilizes cutting-edge deep neural nets, especially recurrent neural nets as the basis for its solution. It is cool that a small company with less than 30 full-time equivalents (FTEs) is able to surpass the translation capabilities of giants such as Google and Baidu. The major driver of why it is better for some language pairs is the quality of its training data. The training data is derived from a significant community of professional translators and other high-quality curated data sources, such as webpages, and documents from the European Parliament, and a curated body of international patents. In a recent survey, DeepL was evaluated as providing the best available machine translation results for 12 language pairs, beaten only by Google in quantity, which provides best-in-class performance in 20 other language pairs. The language translations for the major European languages in particular seem to be owned by DeepL for the second consecutive year. Several other sources are confirmation of those results.²

Challenges: Unfortunately, the website DeepL.com currently only offers its translation services in seven European languages: Dutch, English, French, German, Italian, Polish and Spanish. It misses out on major other languages such as Arabic, Russian, Hindi, Mandarin, Japanese and Portuguese. Also, its translations often have been considered to represent formal language, which can be unfitting if the source text is more colloquial. If DeepL could at least provide an on-premises solution, this could give the company another unique advantage, as many other vendors, such as Google and Baidu, are still sworn into providing just in-the-cloud capabilities.

Who Should Care: All international companies dealing with translations in the above-mentioned seven languages can derive instant benefits and use the free-of-charge online tooling. Also, professional translators are believed to be increasing their own productivity significantly from the high-quality suggestions of DeepL.com. In general, the DeepL story serves as yet another milestone story for all organizations in the world about the importance of high quality data and how even smaller players still can beat IT megavendors.

Phrasee

London, United Kingdom (<https://phrasee.co>)

Analysis by Adrian Lee

Why Cool: Phrasee is cool because it is a robowriter that delivers “human-sounding” email subject lines from a multichannel and contextually aware email marketing platform for advertisers. Phrasee incorporates an in-house-developed natural language generation (NLG) system to generate email subject lines and utilizes a deep learning engine to optimize conversion and open rates for email, Facebook ad campaigns and mobile push notifications. It incorporates natural language processing elements of sentiment analysis, semantic categories of speech (urgency, directness, offbeat, familiarity and curiosity), even emojis, among other linguistic parameters to predict marketing copy performance.

Traditional advertising agencies still rely on human copywriters to generate marketing copy. In digital marketing communications, email marketing is the one of most mature “personalized” channels of digital marketing. In “How Email Marketing Stacks Up in Marketing Leaders’ Spending Priorities,” nearly half of marketing leaders (45%) say they intend to increase email marketing spending in 2017. As businesses increase spending, metrics such as click-through, open and conversion rates are more crucial than ever to the success of email marketing campaigns. Phrasee can autogenerate (theoretically, infinitely) many email subject lines, in accordance with a custom brand-specific lingo and terms corpora. It follows by running agile experiments via split testing (as opposed to normal A/B testing methodology) on email subject lines with randomly selected audiences. After the test period, the successful email subject line is then launched to a wider audience. This adds to its deep learning capability to optimize future campaigns for the advertiser.

Phrasee reports increased open and conversion rates for email campaigns done for clients such as Gumtree, Virgin Holidays and Euromoney. It provides integrations for its technology to work with platforms such as Salesforce, Emarsys and IBM Watson.

Challenges: One challenge will emanate from advertising agencies that will still claim to write better with human copywriters for long-form marketing content. The second challenge will come from conversational platform vendors of NLG systems that are constantly evolving their capabilities with vertical foci to serve industry better. Phrasee is capable of generating shorter-form content (i.e., email subject lines) and optimizing for email campaigns at this point in time. The torrent of digital content production has increased through the deployment of robowriters in content automation platforms by marketers and agencies. Conversational application vendors, as well, are furthering deep learning expertise with training data gathered from multiple input channels (customer service, digital commerce, CRM, etc.) for NLG systems.

As natural language generation improves, the evolution of systems with advanced machine learning techniques that can vary syntax, vocabularies and writing styles, will challenge extant email marketing service providers like Phrasee.

Who Should Care: Digital marketing leaders of consumer organizations who are reliant on digital communications and are constantly under pressure to increase marketing performance and customer engagement should care about Phrasee. General managers of conversational platform vendors interested in boosting short-form text content NLG capabilities to complement their existing capabilities as part of a larger NLP strategy should also care about Phrasee. Product management leaders of multichannel marketing hubs seeking third-party integrations to enrich user experiences and engagement should also care about Phrasee.

Pindrop

Atlanta, Georgia (www.pindrop.com)

Analysis by Jessica Ekholm

Why Cool: Pindrop is cool because it is one of the first companies on the market offering an end-to-end deep neural network (DNN)-based voice biometric engine for cross-channel authentication and fraud detection. The company originated as a voice authentication solution provider for call centers, and it is now aiming to reach beyond its traditional market and into the ever-growing market of voice-enabled devices.

Pindrop's Deep Voice biometric engine is passive, has no dependency on what the speaker is saying, and identifies legitimate and fraudulent speakers with as little as 2 seconds of speech. Data collected from contact center environments over several years has helped Pindrop create the foundation for its DNN-based models for acoustical environments. By being robust to channel variance and background noise, Pindrop's voice solutions works both for contact center, as well as voice-enabled products. Consumers can enroll their voice on a mobile device and subsequently interface with a voice-enabled devices in noisy conditions.

Its Deep Voice engine offers measures against various voice-based attacks, including replay attacks, synthesized voices such as Google Duplex and Lyrebird, as well as voice distortion attacks. In terms of synthesized speech, the biometric engine detects unnatural artifacts produced in computer-generated speech, including advanced text-to-speech synthesizers such as Google WaveNet and Tacotron (as heard in the Google Duplex demo).

The market for voice-enabled devices such as Google Home and Amazon Echo is expected to grow. By 2021, end-user spending on virtual personal assistant (VPA) speaker hardware will reach \$3.52 billion, growing at a compound annual growth rate (CAGR) of 37.5% from 2016 through 2021 (see "Forecast Snapshot: VPA-Enabled Wireless Speakers, Worldwide, 2016-2021"). Thus the use and demand for voice biometrics are expected to rise, especially for purposes such as e-commerce, parental control purposes and safety reasons. Precise voice biometrics will be needed. However, at present there are few contenders in this space that have been able to fully master the art of it.

Challenges: The main challenge for Pindrop is the potential of providers such as Amazon and Google building their own end-to-end DNN voice biometric solution, which could reduce a smaller company's market potential like Pindrop. Another challenge, as with any other market player, is increasing direct competition with other voice ID and voice biometrics providers starting to offer a similar solution. Potential providers include Nuance and Sensory. Competition could also include Daon, which offers multifactor biometric authentication on devices. Pindrop needs to quickly capitalize on its early market opportunity and focus on reach — or rather getting the market to hear about and understand its proposition — as soon as possible to not lose its early mover competitive strength.

Who Should Care: IT leaders in charge of customer experiences and customer care scoping for companies offering voice biometrics solutions should consider Pindrop. These IT leaders can ultimately use these technologies to recognize customers based on voice, device and behavior to

increase the customer experience and to help reduce fraud. Also, providers of voice-enabled devices wanting to integrate voice biometrics into their product suite to offer increased e-commerce and parental control abilities should consider Pindrop. Automotive manufacturers and home security providers looking to deliver innovative and secure voice experiences for their customers should consider Pindrop.

SavantX

Jackson, Wyoming (www.savantx.com)

Analysis by Erick Brethenoux

Why Cool: SavantX provides organizations (today mostly in the utilities industry) with an innovative and powerful natural language processing (NLP) approach to synthesize a vast number of technical documents into an effective searchable and reasoning-prone knowledge representation mechanism.

SavantX leverages deep learning techniques (i.e., nonsymbolic NLP) to process natural language from a remarkably wide source of document formats, a large amount of unstructured, noisy heterogeneous and complex types (including various databases formats). Because the system does not treat those sources through a semantic filter (i.e., it does not categorize the “meaning” of the terms it reads), it can absorb (and correlate) a large amount of terms without being tied (a priori) to a strict vocabulary.

If document corpuses are processed sequentially, through a time dimension for example, the technology can also capture relationships of terms over time through various degrees of abstraction. That capability facilitated by the internal representation that the algorithm leverages allows for the identification of causality links in the text to get to better disambiguation and richer context understanding — for example, the possibility to derive rules and associations.

The underlying graph theory approach is the foundation of the system’s flexibility for an inventive search and visualization capability. A knowledge graph can be viewed from multiple perspectives, each representing a specific understanding of the corpus; in other words, different users with very different intents can leverage that representation from their own understanding of the data. The graph can also be traversed to derive rules and generate inferences.

SavantX has built a productive 3D interactive user interface to navigate the underlying knowledge representation; users navigate the graph using a dynamic model that visualizes clustering of related terms and data. This interface affords the user a rich toolset used to select, subselect and compare data relationships through a 3D collapsible tree-like visualization of the most relevant terms, concepts, key phrases and identifiers.

Challenges: At this stage, SavantX’s chief challenge will be to demonstrate the generalization of the concept to a wider number of industries and business problems. While there is nothing that would, a priori, prevent that generalization, the company will have to carefully apply its still-limited resources for that purpose. In other words, it needs to carefully select its few next areas of implementation.

Nonsymbolic NLP is quickly gaining in popularity through the emancipation of deep neural network (DNN) research and applications. It is also leveraging deep learning to forgotten but powerful probabilistic reasoning techniques such as Bayesian nets, which is an important differentiating approach that SavantX should market judiciously.

Who Should Care: Much of organizations' knowledge remains locked up in unreadable document types and formats. SavantX provides a solution to unlock that potential while providing users with compelling knowledge navigation (and inference) capabilities. Utilities companies can benefit from the company's early industry experience. However, other organizations with difficult-to-read historical records (such as healthcare, manufacturing and government) should explore SavantX's capabilities to open up a new range of critical data in their digitally enabled business environments.

Where Are They Now?

This is a new Cool Vendors report. Next year, we will track the progress of the vendors above.

Gartner Recommended Reading

Some documents may not be available as part of your current Gartner subscription.

“Clarify Strategy and Tactics for Artificial Intelligence by Separating Training and Machine Learning”

“Predicts 2018: Artificial Intelligence”

“Hype Hurts: Steering Clear of Dangerous AI Myths”

“Competitive Landscape: Speech-to-Text Applications”

“Market Guide for Conversational Platforms”

“Market Insight: Creative AI — Assisted and Generative Content Creation”

Evidence

¹ [“DeepL: A Modern-Day David and Goliath Story”](#)

² Only one of the many [discussions about DeepL](#) on ProZ.com, the major social community for professional translators.

[“State of the Machine Translation by Intento.”](#)

[“DeepL: A Modern-Day David and Goliath Story.”](#)

This document is published in the following Market Insights:

- Computing Hardware Worldwide
- Infrastructure Support Services Worldwide
- Mobile Devices Worldwide
- Printing Hardware Worldwide
- Security Solutions Worldwide
- Semiconductor Applications Worldwide
- Software Applications Worldwide

GARTNER HEADQUARTERS**Corporate Headquarters**

56 Top Gallant Road
Stamford, CT 06902-7700
USA
+1 203 964 0096

Regional Headquarters

AUSTRALIA
BRAZIL
JAPAN
UNITED KINGDOM

For a complete list of worldwide locations,
visit <http://www.gartner.com/technology/about.jsp>

© 2018 Gartner, Inc. and/or its affiliates. All rights reserved. Gartner is a registered trademark of Gartner, Inc. and its affiliates. This publication may not be reproduced or distributed in any form without Gartner's prior written permission. It consists of the opinions of Gartner's research organization, which should not be construed as statements of fact. While the information contained in this publication has been obtained from sources believed to be reliable, Gartner disclaims all warranties as to the accuracy, completeness or adequacy of such information. Although Gartner research may address legal and financial issues, Gartner does not provide legal or investment advice and its research should not be construed or used as such. Your access and use of this publication are governed by [Gartner Usage Policy](#). Gartner prides itself on its reputation for independence and objectivity. Its research is produced independently by its research organization without input or influence from any third party. For further information, see "[Guiding Principles on Independence and Objectivity](#)."