

福祉・介護分野における用語と 記録の解析に関する基礎的検討 — サービス提供記録解析の方法と技術 —

生田 正幸*

福祉・介護分野における記録は、多大な労力とコストをかけて作成されてきたにもかかわらず、その多くが限定的な利用にとどまり、情報としての活用が組織的・日常的に行われることは稀であった。しかし、福祉・介護をめぐる状況が大きく変化し厳しさを増すなかで、サービス利用者の生活の質を高め、効果的で効率的なサービス提供を推し進めるために、積極的に記録を活用しようとする動きが強まりつつある。そこで、記録の活用を進めるために、介護分野におけるサービス提供記録を大量に収集し、用語の使われ方に注目したコンピュータ解析により記録の実態を定量的にあきらかにすることで福祉・介護分野における記録活用の可能性を探るとともに、解析のための方法と技術、さらには課題について考察を行った。

キーワード：福祉，介護，福祉情報，サービス提供記録，コンピュータ，形態素解析，品詞，用語の標準化，表記の揺れ，同義語

1. 問題意識

1.1 福祉・介護分野における用語をめぐる課題

福祉・介護は、比較的新しい分野であり、保健・医療・教育など隣接あるいは関連する分野の影響を強く受けていることもあって、使用される「用語」が非常に多彩である。また、概念の一般化していない用語や外来語、独自の略語、地域や個々の施設・機関・組織等において独自に使われている用語など、普遍性を持たな

い用語も少なくない。そのため、ある施設・機関・組織で使用されている用語が、他の施設・機関・組織でまったく使われていなかったり、意味するところが違ったり、正確に理解できなかったりするという問題が生じる。とくに、サービス提供にあたって作成される計画や指示・記録などについて、こうした問題が生じやすく、施設・機関・組織等を超えた情報の伝達あるいは共有・活用に際して障壁となる可能性が大きい。

また、福祉・介護サービスを支援する情報システムの導入が進みつつあることから、情報システムのあり方とも絡んで問題の拡大が懸念される。一般に、福祉・介護サービスに関する情

*立命館大学産業社会学部教授

報システムは、類型化可能なデータを管理する際には分類コードを用い、サービス提供記録など個別性が高いものについては自由記述方式やパターン化した文章を選択させる方法をとっている。情報システムの常として、入力されたデータは、情報ネットワークを通じた伝達や共有、データベース化による蓄積、データ解析による傾向や特徴の分析など、情報としての二次的な活用を容易に行うことができ、サービスの質や効率の向上に活用することが容易な特性を備えている。

ところが、分類コードは情報システム毎に異なり、サービス提供記録などに用いられる用語についてもユーザーの裁量に任せられている場合が多いため、同種の業務を行っているにもかかわらず、システムやユーザーが異なるとデータを構成している分類コードや用語が異なるデータとなってしまう、システム間や地域、施設・機関・組織の間におけるデータの互換性が失われてしまう。その結果、活用することが容易な特性を備えているデータであるにもかかわらず、データの共有や活用が困難という状況に陥ってしまうことになる。

このように、福祉・介護分野においては、用語が情報の伝達や共有・活用のあり方を規定する重要な要素として浮上しつつあり、とりわけ情報システムをめぐる状況においてリスクが拡大している。

1.2 福祉・介護分野における記録をめぐる課題

福祉・介護分野における記録は、様々な形で存在しその役割も多様である。しかし、実態を踏まえるならば、最も重要な役割は、公的サービスである福祉・介護サービスを実施した「証

拠」としての役割であろう。福祉・介護サービス関係の各種の法令が制度の一環として記録のあり方を規定し作成を求めており、相談援助の過程において作成される経過記録やケース記録、入所サービスや居宅サービスの提供にともなって作成される処遇記録やサービス提供記録など、福祉・介護分野において作成されている記録の大半はこの種のものである。

これらの記録は、「証拠」であると同時に、福祉・介護の専門職であるケースワーカーやケアワーカーが、利用者の抱える問題の改善あるいは解決を図り、生活の質の向上を図るために、利用者の問題状況やニーズ、生活状態やサービス利用状況等を、専門的な問題意識と視点から記録したものでもある。ある意味では、サービス利用者の状態やニーズを本人に代わって代弁したものといってもよいだろう。

つまり、福祉・介護分野における記録の多くは、制度の枠組みのなかに位置づけられた「証拠」という役割と、ケースワーカーやケアワーカーによる利用者の状態やニーズの記録あるいは代弁という役割を重層的に担っており、サービスを利用する側にとっても、サービスを提供する側にとっても、大きな価値を持っている。しかし、そうした役割をフルに発揮し、価値を十分に引き出すためには、大きな障壁が存在している。

経済産業省の資金を得て保健医療福祉情報システム工業会（JAHIS）をはじめとする介護サービス実施記録システム普及コンソーシアムが実施した「平成16年度先導的分野戦略的情報化推進事業 介護サービス実施記録の電子的共有を可能とする情報システム基盤に関する調査研究報告書」によれば、介護サービス事業所における「記録の作成手段としては、手書きまたは

ワープロとする回答が7割以上を占め、システムを使用している場合も手書きとの併用が多く、情報システムだけで記録を作成しているケースは全体の2%に過ぎない¹⁾。」との調査結果がある。つまり、用紙に手書きあるいはワープロにより作成されている記録が圧倒的に多い。また、同調査は、記録の記述方法について、「全て文章で記述する方式をとっている割合が約24%、文章による記述を選択型等で補う方式が約40%、選択式等を文章によって補う方式が約33%と大きく回答が分かれており、記録方法が事業所により多様であることがわかる²⁾。」としており、自由記述方式への依存度が大きいことを示している。

つまり、福祉・介護分野における記録は、作成の仕方が手書きであるため、まったく電子化されていないか、ワープロの文書ファイル化にとどまっているため、情報としての利用可能性がきわめて限られている。また、自由記述への依存が強いと記述のあり方も一定せず、先に指摘した分類コードや用語の問題とも相まって、情報としての蓄積や検索、類型化などが難しく、積極的な活用を図ろうとする際の大きな障壁となっている。福祉・介護分野における記録は、サービスを利用する側にとっても、サービスを提供する側にとっても、大きな価値を持つにもかかわらず、情報として活用を図る環境が整っていないのである。

1.3 福祉・介護分野における記録と用語のあり方をめぐる問題意識

2000年における介護保険制度の導入を契機として、福祉・介護サービスの利用のあり方が契約を基礎とする形態に変化するとともに、サービス供給主体が社会福祉法人や企業、NPOな

ど多様化し、複数のサービス提供事業所が一人のサービス利用者に関与するという状況が一般化してきた。

こうした状況において、サービス提供にともなう適正さと透明性を確保するためにサービス実施記録を重視しようとする動きが行政サイドにおいて顕在化する一方、サービスを効果的かつ効率的に提供するためには、個々の利用者に関する計画や指示、記録などの情報を、関係する複数の事業所が共有する必要があるとの認識も進んだ。また、利用者本位の原則を踏まえたサービスを提供するためには、記録の積極的な活用が必要であり、その手段として情報システム（=コンピュータ）の活用が必須であるとの認識も深まりつつある。

これまで福祉・介護分野における記録は、多大な労力とコストをかけて作成されてきたにもかかわらず、その多くが、業務上の申し送りなど日々の備忘や監査指導の際の証拠資料など限定的な利用にとどまり、情報としての活用が組織的・日常的に行われることは稀であった。どちらかといえばルーティンワークの一環として作成され保存され、活用されることもなく倉庫の片隅に眠る記録が多かったのである。

しかし、福祉・介護をめぐる状況が大きく変化し厳しさを増すなかで、サービス利用者の生活の質を高め、効果的で効率的なサービス提供を推し進めるために、積極的に記録を活用しようとする動きが強まりつつある。記録が、情報として大きな価値を持つ資源であることが改めて認識されたといつてよいだろう。

とはいえ、記録の活用を進めるためには、障壁を乗り越え環境を整備しなければならない。まず、必要となるのが記録の電子化である。先ほども述べたように、福祉・介護分野では、未

だに記録の大半が手書きないしはワープロによる作成の状態にとどまっている。紙に書かれた記録、あるいは印刷された記録は、閲覧しやすく利便性が高いように思われがちであるが、蓄積されればされるほど該当するページや部分を見つけることが困難になる。作成されて数週あるいは数ヶ月程度であれば、記憶を頼りに時系列で探し出すこともできるだろうが、数年となると、インデックスの整備などよほど管理が行き届いていない限り速やかに探し出すことはできない。もちろん、手書きやワープロによる記録にも優れた点が数多くあることは確かである。しかし、蓄積や検索・分析をはじめ情報として活用するには電子化されている方が圧倒的に有利であり、その促進が焦眉の課題となっている。

次に、記録の様式や分類コード、使用する用語などについて標準化が必要である。法令が制度の一環として記録のあり方を規定している場合は、ある程度標準化されているが、それでも施設・機関ごとに固有のアレンジや公私の作成ルールが存在している場合が多い。とくに用語については、自由記述方式への依存が高いこともあって、地域毎の方言や独特の言い回し、施設・機関の歴史やスタッフの経歴などに規定された独自の略語や言葉遣いが用いられやすく、同じ意味や内容であるにもかかわらず、施設や機関、時にはスタッフによって用語や表記が異なっている場合が多い。しかし、福祉・介護サービスは、生活全般を援助対象とするため地域差や個人差が非常に大きく、その記録に用いられる用語についてもバラエティに富んでいるのがあるべき姿といえる。そのため用語を一本化し特定の用語に限るのではなく、バラエティに富んだ用語の体系化を図り、ある程度の幅を持たせた標準的な福祉・介護の用語群を構築する

方向が望ましいといえる。

記録の電子化と標準的な用語群の構築が実現すると、福祉・介護分野における記録の共有や活用は大きく前進する。電子化された記録は、情報ネットワークを通じた共有やデータベースとしての蓄積や検索はもちろんのこと、さまざまな分析や分類を高速かつ大量に行うことができる。たとえば、一人の利用者の記録を地域内の複数のサービス提供事業所が共同で作成・活用したり、ケアマネージャが担当ケースの居宅サービス利用状況をリアルタイムで把握・モニタリングしたり、記録そのものの評価や分析を行ったり、記録の傾向や特徴から施設・機関・地域におけるサービス提供の特徴を把握したり、施設利用者の記録をもとに長期的な視点からサービス効果測定や評価を行ったり、蓄積された大量のケース記録を類型化し類似する事例を抽出したりするなどといったことが可能になる。

もちろん、記録をこのように活用するためには、個人情報保護のための体制の確立、異なる情報システムの連携、地域情報ネットワークのあり方、情報セキュリティの整備など、課題が山積している。また、記録を情報として活用するための方法と技術の開発も進んでいないのが実情である。

本稿では、こうした状況と問題意識を踏まえ、福祉・介護分野における記録を情報としていかに活用しうるのか、その可能性について考察することにした。先述したように、こうした取り組みは緒についたばかりであり先行研究も数少ない。そこで、記録の解析により用語の使用状況をあきらかにし、用語をめぐる問題点を把握することによって、記録を情報として活用するための方法と技術について検討することにした³⁾。

2 記録解析と用語体系化のための方法論

2.1 基本的な考え方

一口に福祉・介護分野における用語といっても、その範ちゅうは非常に広い。ここで取り上げようとしている福祉・介護の現場において使用されている用語だけでなく、行政機関や教育機関、研究の場などで用いられている用語があり、より身近な生活場面で用いられている用語もある。福祉・介護は生活に深く関わっており、保健・医療・労働・教育など関係する分野も多いだけに、カバーしている範囲が非常に広く使用されている用語も多彩である。

こうした福祉・介護分野の用語を収集し体系化しようとするとき、その方法は大きく二つに分かれる。第一は、福祉・介護及び関連分野に関する理論的な枠組みを踏まえて用語を選出し体系を構築する方法である。この場合の利点は、用語を体系的に網羅できるという点にあるが、構築にあたって分野全般をカバーする広範な見識が必要であり、それを満たすために作業に携わる人や組織を増やせば増やすほど語彙選定基準があいまいになったり恣意的になり易いという問題がある。また、使用頻度が低く実際にはあまり用いられていないような用語まで収録してしまいがちであること、全体として語数が多くなる傾向にあり使用上の負担が大きいことなども特徴といえる⁴⁾。

第二は、用語使用の実態にもとづき使用頻度の高い語彙を選別して用語群を構築する方法である。この場合の利点は、より実用的な用語群を構築できるという点にあるが、用語を選び出すためのフィールドとなる文書（テキスト）データの確保と用語の切り出し・頻度の集計など

作成作業が容易ではないという問題がある。まず、テキストデータについては、構築しようとする用語群の性質に対応する一定の目的のもとに作成されたもので、用語の出現に極端な偏りのないデータであることが望ましく量的な規模も必要である。また、テキストデータから用語を切り出す作業は煩雑であり、データの規模が大きくなるほど加速度的に負担が重くなるため、手作業による実施はきわめて困難である⁵⁾。

しかし、こうした作業面の問題については、コンピュータを使用し電子化された福祉・介護関係文書の解析を行うことで対応できる。これは、ある意味でコーパス言語学的なアプローチといえる。コーパス (corpus) とは、コンピュータによる検索や分析を前提とする電子化されたテキストデータの大規模な集合を指すが、これを活用して言葉の用い方や言葉同士の結びつきなど言葉の分析・研究を行うのがコーパス言語学である。近年、福祉・介護分野においても、情報化の進展にともなって、サービス提供記録の作成・運用を支援する情報システムが進歩しつつある。先進的な施設・機関等ではこうした情報システムを用いて記録の作成・管理を行っており、とくに介護分野における取り組みが進んでいる。そこで、電子化された介護サービス提供記録を解析して用語を抽出し、用語の使用実態の解明、さらには用語標準化のための体系化への取り組みを進めることにした。

2.2 作業手順

まず問題となるのは、文書（テキスト）データから、いかに用語を切り出すのかという点である。すでに述べたように、テキストデータから用語を切り出す作業は煩雑であり、手作業の場合には作業者が文書を読みながら切り出すべ

き用語を選別し記録，さらにその集計を行うという手順が必要になる。しかも「活用語」である動詞，形容詞，形容動詞，助動詞には活用形が存在するためその部分への対応や，口語体と文語体の調整など，福祉・介護分野に関する知識だけではなく日本語文法の知識も必要となるため作業も限られる。そこで，用語の切り出しにあたっては「形態素解析」の技術を用いて語彙の切り出しを行うこととした。

詳しくは後述するが，形態素とは「それ以上短くすると文脈上意味を持たなくなってしまう最小の言語単位」のことであり，形態素解析とは文を形態素に分解する技術である。形態素解析については，「JUMAN」（京都大学：<http://www.kc.t.u-tokyo.ac.jp/nl-resource/juman.html>），「茶筌」（奈良先端科学技術大学院大学：<http://chasen.naist.jp/hiki/ChaSen/>）などの解析用ソフトウェアが公開されている。今回のプロジェクトにおいては「茶筌」を使用することにした。

また，形態素解析においては，解析のために辞書が必要となる。「茶筌」は，情報処理振興事業協会（IPA）で設定されたIPA品詞体系（THiMCO97）に基づいて一部修正を加えた日本語辞書（ipadic2.7.0）を使用しているが，福祉・介護分野の用語などについては十分対応していないため，先行的に行った作業により得られた語群をユーザー辞書として追加し対応した。

次に，介護サービス提供記録については，できる限り広い範囲から収集することをめざし，居宅サービス及び施設サービスにおけるサービス提供記録に加え，機関間でやり取りされるサービス提供に関わる連絡記録についても，各法人・施設・機関の協力を得て収集し，解析対象データとして整備した。

さらにこれらの形態素解析により切り出した

語彙については，必要な処理を加えたうえで，その出現状況，品詞別の分布などを分析し，用語の使用状況を把握するとともに，分析結果を踏まえつつ，標準的な福祉・介護用語群のあり方について検討を行うこととした。

3 介護サービス提供記録の解析

3.1 解析対象とした介護サービス提供記録の概要

表1に示すように，北海道・愛知県・佐賀県に所在する3つの社会福祉法人に依頼し，高齢者介護を中心とする入所施設および居宅サービス機関において作成された9種類の介護サービス提供記録の提供を受けた。いずれも対象期間は，2003年10月1日から2004年9月30日までの一年間である。

データの詳細については，以下の通り。

○愛知県A園

法人の運営する10種類12カ所の事業所（特別養護老人ホーム，訪問介護〔2カ所〕，デイサービスセンター，居宅介護支援〔2カ所〕，在宅介護支援センター，ケアハウス，老人保健施設，訪問看護ステーション，デイケアセンター，訪問リハビリテーション）の業務連絡用システムに収録されている利用者共通情報データより自由記述部分〔自由記載データ〕を抽出。

○北海道B園

サービス提供記録の作成管理に使用している支援情報システムのデータより介護記録に関する自由記述部分〔介護記録〕を抽出。

○佐賀県C園

法人の運営している7種類7カ所の事業所（特別養護老人ホーム，訪問介護，デイサービスセンター，訪問看護ステーション，デイケア

表1 収集した介護サービス提供記録の件数と有効件数

施設・機関名	収集した自由記述記録	
	件数	有効件数
愛知県A園（法人内連絡記録）	20,250	20,124
北海道B園（特別養護老人ホーム）	15,301	15,301
佐賀県C園（デイケア）	5,797	5,797
佐賀県C園（デイサービス）	10,055	10,055
佐賀県C園（特別養護老人ホーム）	37,187	37,187
佐賀県C園（訪問介護）	9,963	7,852
佐賀県C園（訪問看護）	2,740	1,136
佐賀県C園（養護老人ホーム）	14,813	14,813
佐賀県C園（老人保健施設）	28,942	28,942
計	145,048	141,207

※有効件数とは、収集した記録のうち、意味のある形態素を含む記録の件数

センター、老人保健施設、養護老人ホーム）のそれぞれにおいて記録管理に使用している情報システムのデータより自由記述部分〔申し送り事項〕を抽出。

なお、上記のデータの提供を受けるにあたっては、個人情報保護に配慮し、いずれの記録データについてもデータ項目として利用者の個人名は含まれていない。また、自由記述部分に利用者の個人名と思われる記載が含まれていた場合は解析の過程で当該部分を削除した。

収集した記録データは、すべてテキストデータに変換し、自由記述された記録データの羅列とした（表2）。

3.2 用語の抽出

基本的な考え方

収集した記録データは自由記述文であるため、福祉・介護用語を抽出するには、記録文から1つひとつの語彙を切り出す作業（単語の分かち書き：word segmentation）が必要となる。

この作業は、コンピュータによる形態素解析と呼ばれる技術を用いて処理した。

また、切り出された語彙には、記号や数字、さらには時制や活用によって語尾変化した状態のものも含まれているため、前者については削除処理、後者については、活用・変化を取り払った上で集約を図る必要がある。

さらに、語彙の抽出にあたっては、複合語をどう扱うか等の方針を定める必要がある。今回は、原則として複合語は抽出せず、最小単位の単語での抽出を方針としたが、福祉・介護の用語として必要性がきらかな複合語はあらかじめユーザー辞書に登録して抽出の対象とした。

形態素解析

形態素解析（Morphological analysis）は、構文解析、意味解析、文脈理解などにも自然言語処理の基礎となる要素技術で、ワープロなどの仮名漢字変換や音声認識、情報検索、機械翻訳などに応用されており、「言語学では、意味

表2 収集した介護サービス提供記録の例

- ・医師の指示により採血施行。
- ・医師の指示により検尿施行。
- ・医師の指示により心電図施行。
- ・医師の指示により胸部X-P 施行。
- ・茶話会に参加。今月の行事予定を聞きながら、 どちら焼きとお茶を美味しく召し上がっていたとの事です。
- ・体重測定施行。65.2kg（前月比+ 1.65kg）。
- ・面会あり（姉）。
- ・リハビリの為、 受診。
- ・手芸クラブに参加。10月・11月のカレンダーの色塗りを皆さんと一緒に楽しんでいたとの事です。出来には大変満足されたご様子でした。
- ・誕生会に参加。抹茶のムースを召し上がりながら、 余興の民謡を笑みを浮かべながら聞き入っていたとの事です。
- ・医師の往診を一度も受けた事がないとお話しある。看護師に確認したところ、 7日の往診時に診て頂いたとの事。その旨、 ご本人様に伝えてましたが、 数分では診て貰っていないと話されるが、 他のお客様も主と同様である事を話すと渋々納得されています。
- ・足が痺れるとの事で下肢に湿布貼布。
- ・歯が痛いとの事で、 夕食より主食を粥に変更しています。夕食は全量摂取。
- ・歯痛の訴えは聞かれていません。朝食は全量摂取されています。
- ・歯痛の訴え（-）。食事摂取量にも特変見られず。
- ・希望で、 両足底部に湿布貼用しています。歯痛（-）。
- ・左脇腹痛（+）。湿布貼布。
- ・両足首痛の訴えあり。希望で湿布貼布し様子見る。
- ・歯痛（-）左脇腹の痛み（-）
- ・22時以降、 頻回にナースコールあり。不眠・足の痺れに関する訴えありました。0時過ぎ入眠され、 朝まで良眠されています。
- ・歯科往診。歯の咬み合わせが悪いとの事で受診。義歯調整施行。治療は今回で終了。
- ・両下肢の痺れを訴え、 両足底・両足背部湿布貼布。
- ・希望で両足首に湿布貼布。

表3 日本語の活用変化（五段活用動詞の場合）

活用形	未然形	連用形	終止形	連体形	假定形	命令形
語幹	読	読	読	読	読	読
活用	ま	み	む	む	め	め
語尾等	ない	ます		とき	ば	

を担う最小の言語要素（それ以上分割できない語の単位）を形態素（morpheme）と呼ぶ。この形態素を解析する処理、すなわち、自然言語の文中の単語を識別し（tokenization）、その語形変化を解析し（lemmatization, stemming）、品詞を同定する（part-of-speech tagging）処理⁶⁾とされている。

日本語の場合、たとえば五段活用動詞の「読

む」という単語では、表3のように活用が変化

する。形態素とは、このように変化する形態の「素」を指しており、語のなかで変化しない最小単位を指す。したがって形態素は単語に近い存在であるが、単語そのものを指すわけではなく、形態素解析による解析結果は、活用変化を吸収した形態素の形で出力される。また、解析

表4 形態素解析の例（「文献の読み込みを行う」を形態素解析した場合）

分かち書き	よみ	形態素	品 詞			
			名詞	普通名詞	—	—
文献	ぶんけん	文献	名詞	普通名詞	—	—
の	の	の	助詞	格助詞	—	—
読み	よみ	読む	動詞	—	子音動詞マ行	基本連用形
込み	こみ	込む	動詞	—	子音動詞マ行	基本連用形
を	を	を	助詞	格助詞	—	—
行う	おこなう	行う	動詞	—	子音動詞ワ行	基本形

を行うにあたっては、形態素に関する情報を網羅した形態素辞書と形態素に関する文法知識がベースとなる。表4に、「読む」という動詞を含む文を形態素解析した場合の例を示す。

収集した介護サービス記録データの解析

収集した介護サービス記録データの解析に用いた日本語形態素解析システム「茶筌」（ちゃせん）は、1997年2月19日に version 1.0 正式版が奈良先端科学技術大学院大学自然言語処理学講座（松本研究室）からリリースされたフリーの日本語形態素解析器⁷⁾である。従来の形態素解析器 JUMAN⁸⁾ version 2.0 を改良し、大幅に解析速度を向上させたもので、広く自然言語処理研究に資するため無償のソフトウェアとして開発された。開発拠点である奈良先端科学技術大学院大学のある奈良県生駒市高山町が、日本有数の茶筌の産地であることから、この名前がつけられている。国際表記は「ChaSen」である⁹⁾。

抽出にあたっては、「茶筌」の解析用辞書に、福祉・介護分野に適合するようユーザー辞書（1,391語）を追加した上で、9件の介護サービス記録データのそれぞれについて形態素解析を行った。

なお、C園デイサービスについては、「【送迎

往復利用】」「【午前普通入浴】」「【午後普通入浴】」など、C園デイケアについては、「【送迎：往復，入浴：特殊浴】」「（送迎：往復・入浴：普通浴）」「【送迎：往復，入浴：なし】」などといった定型文の見出し語が、各記録文の冒頭に記載されており、解析結果に影響を及ぼす可能性があったため、該当する部分を削除した上で解析も併せて行った。

解析結果は、表5のように出力される。

表5に見られるように、解析結果には、句読点や「EOS」（End Of Sentence = センテンス終）タグなど無用なデータも含まれているため、これらを削除し形態素数のカウントなどを行った上で不要と考えられる品詞（表6）の語彙を削除した。

3.3 解析結果の検討

形態素数

表7に示すように、形態素解析によって切り出された形態素の総数は191万8,501件、形態素の類型は3万0,012件であった。先述した不要語の削除等を行った結果、有効形態素数は88万0,192件、有効形態素の類型は2万8,725種となった。

これらの値から、解析対象とした記録の特徴について検討してみよう。まず、形態素数＝単

表5 形態素解析結果の例

茶話	チャバナシ	茶話	名詞-一般	
会	カイ	会	名詞-接尾-一般	
に	ニ	に	助詞-格助詞-一般	
参加	サンカ	参加	名詞-サ変接続	
。	。	。	記号-句点	
今月	コンゲツ	今月	名詞-副詞可能	
の	ノ	の	助詞-連体化	
行事	ギョウジ	行事	名詞-一般	
予定	ヨテイ	予定	名詞-サ変接続	
を	ヲ	を	助詞-格助詞-一般	
聞き	キキ	聞く	動詞-自立 五段・カ行イ音便	連用形
ながら	ナガラ	ながら	助詞-接続助詞	
、	、	、	記号-読点	
どら	ドラ	どら	名詞-一般	
焼き	ヤキ	焼き	名詞-接尾-一般	
と	ト	と	助詞-並立助詞	
お茶	オチャ	お茶	名詞-一般	
を	ヲ	を	助詞-格助詞-一般	
美味し	オイシ	美味しい	形容詞-自立 形容詞・イ段	ガル接続
そう	ソウ	そう	名詞-接尾-助動詞語幹	
に	ニ	に	助詞-格助詞-一般	
召し上がっ	メシアガッ	召し上がる	動詞-自立 五段・ラ行	連用タ接続
て	テ	て	助詞-接続助詞	
い	イ	いる	動詞-非自立 一段	連用形
一段				
た	タ	た	助動詞 特殊・タ	基本形
と	ト	と	助詞-格助詞-引用	
の	ノ	の	助詞-連体化	

表6 解析結果より削除対象とした品詞

連体詞, 名詞-動詞非自立的, 名詞-代名詞-一般, 名詞-接尾-特殊, 名詞-接尾-助動詞語幹, 名詞-接続詞的, 名詞-数, 名詞-固有名詞-地域-国, 名詞-固有名詞-地域-一般, 名詞-固有名詞-組織, 名詞-固有名詞-人名-名, 名詞-固有名詞-人名-姓, 名詞-固有名詞-人名-一般, 名詞-固有名詞-一般, 名詞-引用文字列, 動詞-接尾, 接頭詞-名詞接続, 接頭詞-数接続, 接頭詞-形容詞接続, 接続詞, 助動詞, 助詞-連体化, 助詞-並立助詞, 助詞-副助詞/並立助詞/終助詞, 助詞-副助詞, 助詞-副詞化, 助詞-特殊, 助詞-接続助詞, 助詞-終助詞, 助詞-係助詞, 助詞-格助詞-連語, 助詞-格助詞-引用, 助詞-格助詞-一般, 形容詞-接尾, 記号-読点, 記号-句点, 記号-括弧閉, 記号-括弧開, 記号-一般, 記号-アルファベット, 感動詞, 未知語の一部

語数とみなすと、ひとつの記録において形態素数が多いほど、その記録文は長いということになる。図1の「1有効記録当たりの形態素数」（ひとつの有効記録に平均何個の形態素が含まれているか）に見られるように、通所型サービスの「C園デイケア」「C園デイサービス」及び

「C園訪問看護」は1記録あたりの形態素数が40以上で、1つひとつの記録文が他より長い傾向にある。逆に「A園」や「C園特養」「C園養護老人ホーム」「C園老人保健施設」は、1記録あたりの形態素数が10以下で記録文が他より短い傾向にあることがわかる。

表7 解析結果の集計（形態素数）

	有効記録件数*1	出現した形態素（総数）				検討対象とした形態素			
		出現した形態素の総数*2	形態素の種類*3	1有効記録あたりの平均形態素数*4	1形態素（種類）あたりの出現頻度*5	検討対象とした形態素の総数*6	形態素の種類*7	1有効記録あたりの平均形態素数*4	1形態素（種類）あたりの出現頻度*5
A園（法人内連絡記録）	20,124	107,249	4,957	5.3	21.6	48,454	4,264	2.4	11.4
B園（特別養護老人ホーム）	15,301	374,013	7,831	24.4	47.8	164,153	6,908	10.7	23.8
C園（デイケア）	5,797	322,670	1,313	55.7	245.8	143,332	1,114	24.7	128.7
C園（デイサービス）	10,055	571,306	1,607	56.8	355.5	273,596	1,382	27.2	198
C園（特別養護老人ホーム）	37,187	253,835	5,150	6.8	49.3	126,582	4,618	3.4	27.4
C園（訪問介護）	7,852	150,063	3,178	19.1	47.2	64,397	2,811	8.2	22.9
C園（訪問看護）	1,136	46,751	2,730	41.2	17.1	26,699	2,380	23.5	11.2
C園（養護老人ホーム）	14,813	7,961	1,064	0.5	7.5	3,669	893	0.2	4.1
C園（老人保健施設）	28,942	84,653	2,182	2.9	38.8	29,310	1,875	1.0	15.6
合計	141,207	1,918,501	30,012	13.6	63.9	880,192	28,725	6.2	30.6

- * 1：有効な文が記入されていないもの、主に空白の記録を除いた件数
- * 2：形態素解析により切り出された形態素の累計。記号、括弧、数字などを含む
- * 3：「* 2」について同一の形態素を1種類として数えた場合の累計
- * 4：値が大きいかほど1記録あたりの記述が長いことを意味する
- * 5：値が大きいかほど同じ用語が用いられている頻度が高く、同じような内容の記述が多いことを意味する
- * 6：表6の品詞など削除した件数
- * 7：「* 6」について同一の形態素を1種類として数えた場合の累計

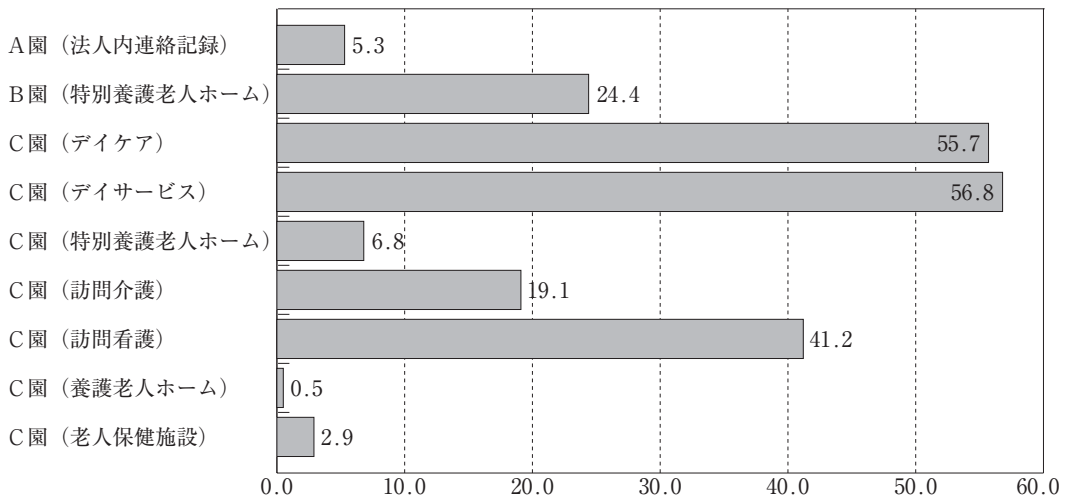


図1 1有効記録あたりの形態素数

また図2の「1形態素あたりの出現頻度」（同じ形態素が当該事業所の解析対象記録データのなかに平均何回出現したか）を見ると、「C園デイケア」「C園デイサービス」における出現頻度が200回あるいは300回を超えており、同じ用語が頻繁に使用されていることがわかる。いいかえれば、同じような記述や定型的な記述が多用されていることになる。これに対し

て、「C園訪問看護」「A園」「C園老人保健施設」は出現頻度が40以下で、多様な用語が使用されており、記録内容が多彩であることがわかる。

このように、今回提供を受けたサービス提供記録については、デイケア・デイサービスにおいて1記録あたりの長さが相対的に長く定型的な記述がなされる傾向にあること、特別養護老人ホームや老人保健施設において1記録あたり

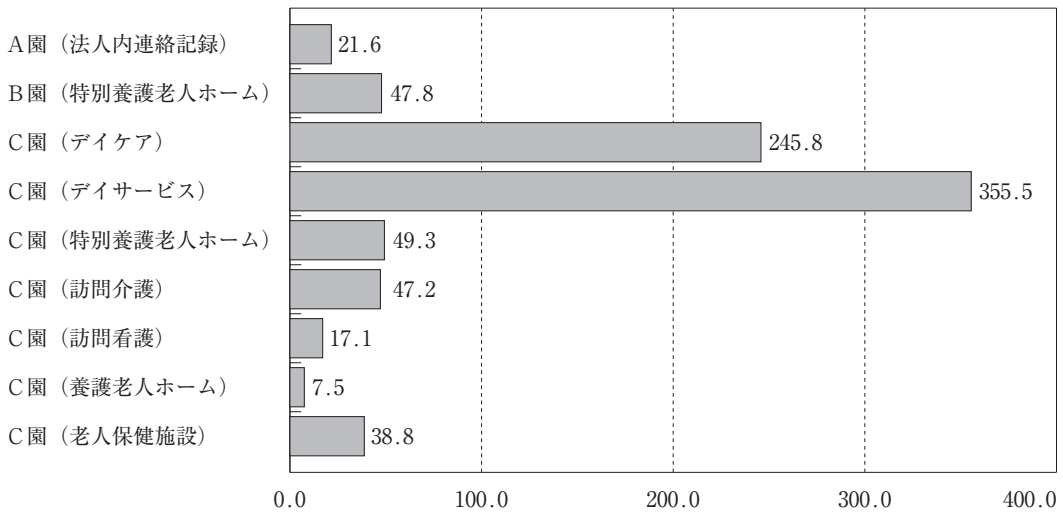


図2 1形態素（種類）あたりの出現頻度

の長さは相対的に短いが多様な内容が記述されていることがあきらかになった。

なお、C園関係者とともにこうした特徴について検討と議論を行ったところ、上に述べたデイケア・デイサービスにおける記録については、サービス実施にともなう制度の規定により利用者（及び家族）に加算対象となるサービス内容を含むサービス実施記録を提供する必要があり、当日のサービス利用状況を逐一記載する関係でどうしても定型的な記録になりがちのことであり、記録の長さについては、通所サービス機関における滞在時間が6時間程度であることから1～2人の担当職員が当日の記録を一気に記述するため長くなることとであった。これに対して、特別養護老人ホームや老人保健施設については、利用者は24時間施設内に滞在しており、職員は交代勤務であるため、気がついたこと、記録すべきことを、その時々記述するため、短く多様な内容の記録になりがちであるとの指摘を得た。

語彙

今回提供を受けたサービス提供記録を形態素解析することで総数88万0,192件の語彙（＝検討対象とした形態素であり、重複する同一語彙もそれぞれ1件とカウントしている）が切り出されたが、表8に見られるように、最も出現回数が高かった上位5位は、①動詞の「さ：“する”の活用形」（44,008回）、②動詞の「い：“いる”の活用形」（35,335回）、③名詞の「参加」（23,340回）、④名詞の「利用」（19,666回）、⑤名詞の「午後」（18,593回）であった。重複する同一語彙を集計した結果、出現した語彙の種類は総数26,245種、うち上位50位までの語彙の出現回数の合計は43万8,985回であり、切り出した語彙の総数88万0192件の49.9%が上位50位までの語彙によって占められていることになる。

品詞別構成として名詞について見ると、総数63万3,364件の名詞が切り出され、そのうち最も出現回数が高かった上位5位は、①「参加」（23,340回）、②「利用」（19,666回）、③「午後」（18,593回）、④「入浴」（18,460回）、⑤「午前」

表8 出現回数の多い語彙（総数：上位50位まで）

順位	分かれ書き	読み	語彙	品詞分類	出現回数
1	さ	サ	する	動詞	44,008
2	い	イ	いる	動詞	35,335
3	参加	サンカ	参加	名詞	23,340
4	利用	リヨウ	利用	名詞	19,666
5	午後	ゴゴ	午後	名詞	18,593
6	入浴	ニュウヨク	入浴	名詞	18,460
7	午前	ゴゼン	午前	名詞	17,357
8	往復	オウフク	往復	名詞	15,937
9	送迎	ソウゲイ	送迎	名詞	15,936
10	中	チュウ	中	名詞	13,894
11	し	スレ	する	動詞	11,167
12	普通	フツウ	普通	名詞	10,327
13	楽しま	タノシマ	楽しむ	動詞	9,891
14	レクリエーション	レクリエーション	レクリエーション	名詞	9,687
15	時	ジ	時	名詞	9,412
16	者	シャ	者	名詞	8,495
17	他	タ	他	名詞	8,257
18	方	ハウ	方	名詞	7,759
19	あり	アリ	ある	動詞	7,475
20	クラブ	クラブ	クラブ	名詞	7,293
21	交流	コウリユウ	交流	名詞	7,282
22	図れ	ハカレ	図れる	動詞	6,671
23	その後	ソノゴ	その後	名詞	5,859
24	行う	オコナウ	行う	動詞	5,556
25	様子	ヨウス	様子	名詞	5,551
26	あり	アリ	あり	動詞	5,139
27	浴	ヨク	浴	名詞	4,927
28	カラオケ	カラオケ	カラオケ	名詞	4,613
29	リハビリ	リハビリ	リハビリ	名詞	4,443
30	する	スル	する	動詞	4,372
31	午前中	ゴゼンチュウ	午前中	名詞	4,329
32	確認	カクニン	確認	名詞	4,212
33	て	テ	てる	動詞	4,196
34	夜間	ヤカン	夜間	名詞	4,048
35	時	ジ	時	名詞	3,980
36	行なっ	オコナッ	行なう	動詞	3,826
37	取り組ま	トリクマ	取り組む	動詞	3,671
38	ゲーム	ゲーム	ゲーム	名詞	3,658
39	℃			未知語	3,648
40	過ぎさ	スゴサ	過ぎす	動詞	3,635
41	塗布	トフ	塗布	名詞	3,606
42	後	ノチ	後	名詞	3,520
43	為	タメ	為	名詞	3,472
44	良眠	リョウミン	良眠	名詞	3,438
45	軟膏	ナンコウ	軟膏	名詞	3,415
46	に	ニ	にる	動詞	3,310
47	休養	キュウヨウ	休養	名詞	3,275
48	日	ニチ	日	名詞	3,126
49	談話	ダンワ	談話	名詞	2,979
50	服薬	フクヤク	服薬	名詞	2,939

(17,357回)であった。出現した名詞の種類は総数16,862種、うち上位50位までの語彙の出現回数の合計は32万1,742回であり、切り出した名詞の総数63万3,364件の50.8%が上位50位までの語彙によって占められていることになる。

動詞については、総数21万4,548件の動詞が切り出され、そのうち最も出現回数が高かった上位5位は、①「さ」（「する」の活用形：44,008回）、②「い」（「いる」の活用形：35,335回）、③「し」（「する」の活用形：18,593回）、④「楽しま」（「楽しむ」の活用形：9,891回）、⑤「あり」（「ある」の活用形：17,357回）であった。出現した動詞の種類は総数6,894種、うち上位50位までの語彙の出現回数の合計は17万5,857回であり、切り出した動詞の総数21万4,548件の82.0%が上位50位までの語彙によって占められていることになる。

形容詞については、総数13,127件の形容詞が切り出され、そのうち最も出現回数が高かった上位5位は、①「なし」（「ない」の活用形：2,923回）、②「なく」（「ない」の活用形：2,496回）、③「ない」（573回）、④「良く」（「良い」の活用形：426回）、⑤「多く」（「多い」の活用形：419回）であった。出現した形容詞の種類は総数1,036種、うち上位50位までの語彙の出現回数の合計は11,121回であり、切り出した形容詞の総数13,127件の84.7%が上位50位までの語彙によって占められていることになる。

このように、今回解析対象としたサービス提供記録においては、全般的に見て使用されている語彙はそれほど多くなく、むしろ偏っていることがあきらかになった。

表記の揺れ・誤り・同義語

なお、語彙について検討するなかで、表記の

揺れや誤りを含む同義語の問題が表面化してきた。例えば、一般に「ケアマネジャー」と称されている職種が、介護保険制度のもとでは「介護支援専門員」と表記され、略して「ケアマネ」と呼ばれる場合もあるように、使う人の立場やその時の状況・文脈によって同義ではあるが異なる言葉（同義語）が使われることが少なくない。また、介護に際して体の向きを変える行為について「体位交換」「体交」「たいこう」「体位変換」など、様々な表記が用いられたり、「床ずれ」「床づれ」「とこずれ」「とこづれ」「褥瘡」「じょくそう」「褥そう」「じょく瘡」「ジョクソウ」,「ベッド」「べっと」など、仮名・漢字の使い分けや送り仮名の違いによる異なった表記、記述に伴う勘違い・誤りなどが生じている場合も多い。

表9に示すように、今回の解析で把握された名詞の種類は1万6862種であるが、事業所間における重複の整理、出現頻度1回の語彙の削除などの処理を行うと4,138種の語彙リストが得られた。これを便宜的に分類すると表10のような構成となり、うち同義語（主に標記の揺れや誤りによるもの）は233語、全体の5.6%であった。同義語の出現比率が多い分類は、「介護用具・機器・用品」（14.6%）,「身体部位」（14.0%）,「人」（12.7%）,「食品・飲料」（12.6%）であり、実数として多かったのは、「身体部位」（36語）,「食品・飲料」（31語）,「時・月日・季節」（26語）,「状態（医療・看護・介護）」（22語）である¹⁰。

このような表記の揺れや記述の誤りなどは、関係者が直接読む場合にはほとんど問題にならず、ほぼ無意識のうちに同義語あるいは正しい表記に訂正・解釈され理解される。しかし、コンピュータがデータとして取り扱う場合には、異なる文字コードとして認識されてしまうため

表9 事業行列に見た語彙の出現状況

■出現頻度（実数）		（単位：種）					
出 所	合計 出現頻度	形容詞	動詞	副詞	未知語	名詞	
A園（法人内連絡記録）	48,454	1,446	9,593	1,063	931	35,421	
B園（特別養護老人ホーム）	164,153	4,630	39,684	2,985	1,920	114,934	
C園（デイケア）	143,332	791	41,063	1,254	751	99,473	
C園（デイサービス）	273,596	720	65,461	818	2,459	204,138	
C園（特別養護老人ホーム）	126,582	2,633	36,390	1,620	843	85,096	
C園（訪問介護）	64,397	782	9,159	708	2,556	51,192	
C園（訪問看護）	26,699	1,528	4,372	313	316	20,170	
C園（養護老人ホーム）	3,669	57	1,016	46	13	2,537	
C園（老人保健施設）	29,310	540	7,810	236	321	20,403	
計	880,192	13,127	214,548	9,043	10,110	633,364	

■出現頻度（構成比）		（単位：種）					
出 所	合計 出現頻度	形容詞	動詞	副詞	未知語	名詞	
A園（法人内連絡記録）	48,454	3.0%	19.8%	2.2%	1.9%	73.1%	
B園（特別養護老人ホーム）	164,153	2.8%	24.2%	1.8%	1.2%	70.0%	
C園（デイケア）	143,332	0.6%	28.6%	0.9%	0.5%	69.4%	
C園（デイサービス）	273,596	0.3%	23.9%	0.3%	0.9%	74.6%	
C園（特別養護老人ホーム）	126,582	2.1%	28.7%	1.3%	0.7%	67.2%	
C園（訪問介護）	64,397	1.2%	14.2%	1.1%	4.0%	79.5%	
C園（訪問看護）	26,699	5.7%	16.4%	1.2%	1.2%	75.5%	
C園（養護老人ホーム）	3,669	1.6%	27.7%	1.3%	0.4%	69.1%	
C園（老人保健施設）	29,310	1.8%	26.6%	0.8%	1.1%	69.6%	
計	880,192	1.5%	24.4%	1.0%	1.1%	72.0%	

■語彙類型（実数）		（単位：種）					
出 所	合計 語彙類型	形容詞	動詞	副詞	未知語	名詞	
A園（法人内連絡記録）	4,264	151	1,079	163	100	2,771	
B園（特別養護老人ホーム）	6,908	297	1,996	257	143	4,215	
C園（デイケア）	1,114	38	216	33	23	804	
C園（デイサービス）	1,382	43	272	36	37	994	
C園（特別養護老人ホーム）	4,618	181	1,508	137	52	2,740	
C園（訪問介護）	2,811	112	720	84	81	1,814	
C園（訪問看護）	2,380	112	444	87	104	1,633	
C園（養護老人ホーム）	893	31	240	21	3	598	
C園（老人保健施設）	1,875	71	419	50	42	1,293	
計（重複を含む）	26,245	1,036	6,894	868	585	16,862	

■語彙類型（構成比）		（単位：種）					
出 所	合計 語彙類型	形容詞	動詞	副詞	未知語	名詞	
A園（法人内連絡記録）	4,264	3.5%	25.3%	3.8%	2.3%	65.0%	
B園（特別養護老人ホーム）	6,908	4.3%	28.9%	3.7%	2.1%	61.0%	
C園（デイケア）	1,114	3.4%	19.4%	3.0%	2.1%	72.2%	
C園（デイサービス）	1,382	3.1%	19.7%	2.6%	2.7%	71.9%	
C園（特別養護老人ホーム）	4,618	3.9%	32.7%	3.0%	1.1%	59.3%	
C園（訪問介護）	2,811	4.0%	25.6%	3.0%	2.9%	64.5%	
C園（訪問看護）	2,380	4.7%	18.7%	3.7%	4.4%	68.6%	
C園（養護老人ホーム）	893	3.5%	26.9%	2.4%	0.3%	67.0%	
C園（老人保健施設）	1,875	3.8%	22.3%	2.7%	2.2%	69.0%	
計	26,245	3.9%	26.3%	3.3%	2.2%	64.2%	

表10 抽出された名詞の構成と同義語の出現状況

分類	用語数			同義語の割合
	総数	見出し語数	同義語数	
病名	38	34	4	10.5%
薬品	50	49	1	2.0%
身体部位	257	221	36	14.0%
状態（身体・コミュニケーション）	104	100	4	3.8%
状態（心理・精神）	117	112	5	4.3%
状態（生活）	463	459	4	0.9%
状態（医療・看護・介護）	475	453	22	4.6%
行為（医療・看護・介護）	205	203	2	1.0%
行為（生活全般）	519	508	11	2.1%
食品・飲料	247	216	31	12.6%
衣類	48	45	3	6.3%
介護用具・機器・用品	48	41	7	14.6%
生活用具・機器・用品	259	243	16	6.2%
生活一般	38	38	0	0.0%
レクリエーション・遊び・趣味	183	172	11	6.0%
人	79	69	10	12.7%
家族・親族・知人・隣人	68	67	1	1.5%
場所・方向	221	204	17	7.7%
制度・施策・社会資源	65	62	3	4.6%
時・月日・季節	276	250	26	9.4%
自然	89	81	8	9.0%
程度	181	175	6	3.3%
単位	80	75	5	6.3%
色彩	28	28	0	—
計	4,138	3,905	233	5.6%

同義語あるいは正しい表記としての扱いができず、同じ内容が記述されているにも関わらず異なった内容として取り扱われてしまう。

こうした事態を回避するためには、表記の揺れや記述の誤りをシソーラス¹¹⁾ という一種の辞書を用いて吸収する方法が有効と考えられる。先の例で言えば「体位交換」「体交」「たい

こう」「体位変換」のいずれを入力しても、標準見出し語である「体位交換」として処理するようシステムを構築するわけである。

品詞類型別構成

表9に見られるように、語彙の出現回数を事業所別・品詞類型別に見ると、全体の72.0%を

名詞が占め、動詞が24.4%、形容詞は1.5%にとどまっている。事業所別に見ると、「C園訪問介護」「C園訪問看護」では、名詞が75%水準を超え、形容詞については、「C園訪問看護」(5.7%)、「A園」(3.0%)、「B園」(2.8%)以外では2%水準を下回り、「C園デイケア」(0.6%)、「C園デイサービス」(0.3%)では、1%を下回る低い水準となっている。

つまり、収集したサービス提供記録は、大半が名詞によって構成されており、形容詞はごく僅かしか使用されていないということがあきらかになった。いいかえれば、「○○を××した。」というようなサービス実施に関する記録文が大半を占め、「うれしい」「楽しい」「苦しい」「悲しい」「辛い」などといった形容詞を用い、サービス利用者の意識や感情、生活などに関する状態や動向等を記録したもの（生活記録）が乏しいことになる。（参考：表11）

また、同じく表9に見られるように、サービス提供記録に使用されている語彙の種類（用語の種類）については、最も多かった「B園」でも6,908種であり、以下「C園特養」(4,618種)、「A園」(4,264種)、「C園訪問介護」(2,811種)、「C園訪問看護」(2,380種)、「C園老健」(1,875種)、「C園デイサービス」(1,382種)、「C園デイケア」(1,114種)、「C園養護」(893種)の順となっている。使われている語彙の種類は、比較的限定されていることがわかる。

また、語彙の大半を占めている名詞について語彙の種類（用語の種類）を見ると、「B園」(4,215種)、「A園」(2,771種)、「C園特養」(2,740種)、「C園訪問介護」(1,814種)、「C園訪問看護」(1,633種)、「C園老健」(1,293種)、「C園デイサービス」(994種)、「C園デイケア」(804種)、「C園養護」(598種)であり、5,000種以内

の範囲に収まっている。

先に指摘したように、今回解析対象としたサービス提供記録においては、使用されている語彙の種類はそれほど多くなく、また大半が名詞であって、日常的に使用されている用語はさらに偏っていることがあきらかになった。

解析結果のまとめ

以上の分析から、今回収集したサービス提供記録に使用されている用語は、それほど多様ではなくむしろ偏っていることがあきらかになった。それゆえ記録そのものについても、その内容的な広がりにはさほどではなく比較的限られていると考えられる。また、語彙の表記に揺れや誤りが多いことや使用されている用語の大半が名詞であり最大でも5,000種の範囲にとどまることもあきらかになった。福祉・介護分野における用語の標準化を図るためには、関係する名詞の標準化が必須であり、表記の揺れや誤りを吸収するためにシソーラス化の形ですすめるべきといえよう。

なお、記録に用いられている語彙（用語）については、それぞれの事業所で用いられている記録作成・運用支援用の情報システムの特徴によって、自由記述部分に記載される内容や様式に差異が生じることもあきらかになった。たとえば、C園で使用されているシステムは、サービス提供内容やバイタルなど基本的な部分については、必要な項目にチェックを入れることで記録が行われるようになっており、チェック項目に該当しない項目について自由記述による記録が作成されている。これに対して、B園で用いられているシステムでは、自由記述による記録の比重が高い。また、自由記述部分について、システムにより記入文字数の制約がなされ

ている場合には、必然的に圧縮した記述となることも留意すべき点でありC園の場合がこれに該当する。

4 サービス提供記録解析をめぐる課題と展望

冒頭で指摘したように、福祉・介護分野における記録の多くは、手書きあるいはワープロで作成されており自由記述への依存も強い。このため記述のあり方が一定せず、情報としての利用可能性もきわめて限られているため、積極的に活用することが難しい。しかし、福祉・介護分野における記録は、サービスを利用する側にとっても、サービスを提供する側にとっても、資源として大きな価値を持っている。そこで、福祉・介護分野における記録を情報としてどのように活用できるのかについて、福祉・介護情報の活用のあり方を探る取り組みのひとつとして検討を行ってきた。

その結果、自由記述形式で作成されたサービス提供記録は電子化されていれば、形態素解析の技術を用いて解析することにより、さまざまな情報を引き出すことができ、記録を定量的に把握・評価することも可能であることがあきらかになった。従来、こうした取り組みは、作業量が膨大であることからほとんど行われておらず、行われていてもごく限られた量の記録を対象とするものであったが、今回の取り組みにより、自由記述形式で作成された大量のサービス提供記録を解析する際の手法について一定の方向性を見いだすことができたと考える。

4.1 サービス提供記録の解析をめぐる課題

しかし、サービス提供記録の解析を巡っては、残された課題も多い。以下に列挙しておく。

①記録の電子化の推進

最も大きな課題は記録の電子化が遅れている点であり、行政の電子化や福祉・介護分野の情報化などに対する政策的な取り組みなど積極的な対応が不可欠である。

②標準化の推進

解析をより適切に行うためにも、福祉・介護分野の記録作成に使用される用語や記録様式などの標準化が必要である。

③より幅広いサービス提供記録の収集・解析

今回行った分析・検討は、対象事業所が限られていたため、集計・解析結果にも協力いただいた事業所の特性や傾向が強く反映されたと考えられる。標準的な用語群とそれらの出現特性モデルを得るには、より幅広いサービス提供記録の収集・解析が必要である。

④複合語の解析

これまでの解析は、「茶筌」による単語切り出しをベースとしているため、複合語の切り出しや解析が不十分である。サービス提供記録では、複合語が重要な要素となっており、改めて複合語に焦点を当てた解析と処理を行う必要がある。

⑤解析プロセスのツール化

多数のサービス提供記録から有効な用語を切り出し整理・解析する手順は、作業プロセスに未知の部分が多く試行錯誤の積み重ねであったこともあり、今回のプロジェクトでは多くを手作業によって行った。記録の解析をより効率的に行い、用語群を用いた記録の検討などを進めるためにも、解析プロセスのツール化を図る必要がある。

サービス提供記録解析の展望

記録の電子化、用語等の標準化、記録解析の

推進により成果が期待される点を以下に列挙しておく。

①記録の共有と活用の可能性の拡大

現状では、用語の不統一や表記の揺れのため、異なる法人間はもちろんのこと、同一法人であっても事業所や担当が異なると、記録の電子的な共有や活用が困難な状況にある。用語の標準化が図られれば、状況が大きく改善され、地域包括ケアや小規模多機能型の推進を図るうえで有効と考えられる。

②記録の評価と検討への活用

用語群をもとにサービス提供記録の解析を行うことにより、用語や品詞の出現状況などによる記録の定量的な評価が可能となり、記録そのものの検討にとどまらず、スタッフの観察や気づきのあり方、サービス提供のあり方についても、より具体的な検討が可能となる。

③処遇記録の検索や類型化への活用

現在のサービス提供記録では、電子化されていても用語の不統一や表記の揺れのため、データを蓄積しても有用な検索・分析ができない点

が問題になっている。標準用語群シソーラスを作成すれば、あいまい検索を含め、的確な処遇記録の検索ができるようになり、類似事例の検索や記録を通したケースの類型化などが可能となる。

④処遇記録作成環境の改善

標準用語のシソーラスを整備し、パソコンの日本語入力環境（IME）と連動させるようになれば、処遇記録の表記の揺れや誤記入を防ぐことができ、処遇記録自体の質の向上につながると期待できる。

⑤処遇記録解析によるサービス改善サイクルの確立

処遇記録の出現単語解析によって、各事業所の業務特性の評価、記録者の記録能力やくせの評価、利用者の状態変化の把握などがそれぞれ可能になる。これらを的確に行う手法を確立すれば、処遇記録の解析結果をもとにサービス改善を図る PDCA（Plan [計画]・Do [実施]・Check [監視]・Action [改善]）サイクルが実現できる。

表11 事業所別にみた語彙の

順位	B園（特別養護老人ホーム）			C園（特別養護老人ホーム）			C園（老人保健施設）			語彙
	語彙	形態素	出現数	語彙	形態素	出現数	語彙	形態素	出現数	
1	い	いる	6,024	さ	する	4,305	さ	する	3,073	さ
2	し	する	5,342	行う	行う	3,538	良眠	良眠	2,511	利用
3	様子	様子	4,803	時	時	3,040	時	時	1,326	い
4	さ	する	2,317	時	時	2,228	午後	午後	481	参加
5	夜間	夜間	2,286	実施	実施	2,035	帯	帯	452	午前
6	あり	ある	2,215	あり	あり	1,980	楽しま	楽しむ	431	午後
7	参加	参加	2,190	に	に	1,959	畑	畑	415	往復
8	BT	BT	2,108	あり	ある	1,915	行う	行う	383	送迎
9	日中	日中	2,049	夜間	夜間	1,633	入浴	入浴	339	中
10	時	時	1,890	する	する	1,432	に	に	327	楽しま
11	なく	ない	1,722	為	為	1,416	あり	ある	306	者
12	施行	施行	1,682	午後	午後	1,395	あり	あり	264	他
13	する	する	1,592	中	中	1,337	夕食	夕食	258	方
14	面会	面会	1,533	介助	介助	1,315	参加	参加	254	交流
15	為	為	1,363	声掛け	声掛け	1,174	田	田	253	図れ
16	事	事	1,175	見	見る	1,161	歌	歌	232	クラブ
17	あり	あり	1,163	日	日	1,059	体温	体温	214	入浴
18	いる	いる	1,139	午前	午前	1,041	経過	経過	214	普通
19	特変	特変	1,044	居室	居室	1,039	夜	夜	213	レクリエ
20	事	事	1,042	オムツ	オムツ	1,021	℃	℃	211	ーション
21	状況	状況	898	交換	交換	997	皆さん	皆さん	209	行なっ
22	時	時	875	夕食	夕食	990	過ごさ	過ごす	209	て
23	排便	排便	845	し	する	975	時	時	208	ゲーム
24	様	様	815	朝食	朝食	849	レク	レク	203	大会
25	良眠	良眠	805	訴え	訴え	843	深夜	深夜	195	服薬
26	過ごし	過ごす	793	休ま	休む	833	中	中	184	確認
27	会	会	758	下	下	819	施行	施行	180	訓練
28	変わり	変わる	725	岸	岸	787	本日	本日	174	歩行
29	特に	特に	707	ホール	ホール	765	する	する	174	塗布
30	受診	受診	699	参加	参加	764	摂取	摂取	173	軟膏

出現状況（全品詞：上位30位まで）

（単位：回）

C 園（デイサービス）		C 園（デイケア）			C 園（訪問介護）			C 園（訪問看護）		
形態素	出現数	語彙	形態素	出現数	語彙	形態素	出現数	語彙	形態素	出現数
する	18,059	さ	する	14,508	℃		2,192	あり	あり	1,309
利用	17,147	い	いる	13,161	準備	準備	2,113	なし	ない	1,148
いる	15,363	入浴	入浴	11,121	確認	確認	1,889	あり	ある	708
参加	14,542	送迎	送迎	5,809	cc	cc	1,622	発赤	発赤	471
午前	14,311	午後	午後	5,726	居室掃除	居室掃除	1,523	排便	排便	400
午後	10,688	往復	往復	5,479	等	等	1,452	症状	症状	281
往復	10,438	参加	参加	5,370	行う	行う	1,435	食思	食思	279
送迎	9,866	浴	浴	4,888	昼食	昼食	1,213	時	時	264
中	9,686	普通	普通	4,877	排尿	排尿	1,139	塗布	塗布	259
楽しむ	7,392	レクリエーション	レクリエーション	4,611	戸締り	戸締り	1,091	昨日	昨日	255
者	7,220	その後	その後	4,363	服薬	服薬	1,051	血糖	血糖	229
他	7,215	午前中	午前中	3,997	塗布	塗布	991	本日	本日	218
方	7,206	リハビリ	リハビリ	3,737	ケア	ケア	978	良好	良好	218
交流	6,925	取り組み	取り組み	3,661	介助	介助	974	こと	こと	215
図れる	6,671	後	後	3,281	尿	尿	952	胸部	胸部	207
クラブ	6,667	休養	休養	2,685	トイレ	トイレ	906	右	右	205
入浴	5,676	し	する	2,588	さ	する	885	訪問	訪問	197
普通	5,378	談話	談話	1,897	軟膏	軟膏	857	部	部	186
レクリエーション	4,860	楽しみ	楽しむ	1,816	朝食	朝食	854	左	左	186
カラオケ	4,199	午前	午前	1,702	尿量	尿量	729	インスリン	インスリン	182
行なう	3,754	中	中	1,689	就寝	就寝	725	排尿	排尿	181
てる	3,549	他者	他者	1,639	モーニング	モーニング	700	中	中	177
ゲーム	2,779	過ぎ	過ぎ	1,556	夕食	夕食	592	軽度	軽度	177
大会	1,765	時	時	1,510	摂取	摂取	570	に	に	162
服薬	1,739	テレビ	テレビ	1,153	事	事	559	し	する	157
確認	1,719	観賞	観賞	1,081	室温	室温	520	内服	内服	151
訓練	1,669	測定	測定	883	し	する	515	臀部	臀部	145
歩行	1,655	入浴後	入浴後	820	て	てる	468	回	回	144
塗布	1,384	再び	再び	745	移乗介助	移乗介助	439	痛み	痛み	143
軟膏	1,367	者	者	609	入浴介助	入浴介助	417	軟膏	軟膏	133

注

- 1) 「平成16年度先導的分野戦略的情報化推進事業 介護サービス実施記録の電子的共有を可能とする情報システム基盤に関する調査研究報告書」(介護サービス実施記録システム普及コンソーシアム:2005年3月)42ページ参照
- 2) 「平成16年度先導的分野戦略的情報化推進事業 介護サービス実施記録の電子的共有を可能とする情報システム基盤に関する調査研究報告書」(介護サービス実施記録システム普及コンソーシアム:2005年3月)46ページ参照
- 3) 本稿は、筆者が、その企画と実施に参画し、データ解析・報告等を担当した平成16年度長寿社会福祉基金(一般分)による助成事業「標準的な福祉介護用語群の開発・活用事業」(財団法人医療情報システム開発センター)の成果に基づいて執筆したものである。
- 4) 経済産業省による次世代介護支援情報システム研究プロジェクトの一環として、筆者も参画して実施した「平成13年度保健医療福祉情報流通推進事業 次世代介護情報システムの開発 次世代介護情報システムの展望及び福祉介護関連用語・コードの標準化作業」および「平成14年度情報経済基盤整備(保健医療福祉分野の標準化に向けたシステム設計・実証研究) 次世代介護支援情報システムに係わる調査研究」における研究成果を踏まえている。
- 5) たとえば、横山正博氏による「社会福祉援助技術における記録の課題」宇部短期大学学術報告・第34号所収(July, 1997)
- 6) テキスト・マイニング研究会による WordMiner サポートサイトに収録されている「テキスト型データのマイニングとその応用」の「第IV部 形態素と分かち書き処理」(株式会社 平和情報センター 保田明夫氏による http://wordminer.comquest.co.jp/wmtips/pdf/H15_01-4.pdf)
- 7) 奈良先端科学技術大学院大学自然言語処理学講座によるホームページ「ChaSen's Wiki」(<http://chasen.naist.jp/hiki/ChaSen/>)より入手することができる。
- 8) 東京大学 大学院情報理工学系研究科 電子情報学専攻 黒橋研究室のホームページ (<http://www.kc.t.u-tokyo.ac.jp/nl-resource/juman.html>)参照
- 9) 奈良先端科学技術大学院大学自然言語処理学講座によるホームページ内の <http://chasen.aist-nara.ac.jp/chasen/whatis.html> を参照
- 10) 同義語チェックについては、浅井学園大学人間福祉学部講師小沼春日氏の協力による。
- 11) 「シソーラス」(thesaurus)とは、1つの見出し語について、同義語、広義語、狭義語、関連語等と関連づけ分類整理した用語集のこと。

参考文献

- 「宇部短期大学学術報告」第34号, July, 1997
- 「平成13年度保健医療福祉情報流通推進事業 次世代介護情報システムの開発 次世代介護情報システムの展望及び福祉介護関連用語・コードの標準化作業報告書」(財団法人医療情報システム開発センター 2002年3月)
- 「平成14年度 情報経済基盤整備(保健医療福祉分野の標準化に向けたシステム設計・実証研究) 次世代介護支援情報システムに係わる調査研究」成果報告書(アライド・ブレインズ株式会社 2003年3月)
- 「平成16年度先導的分野戦略的情報化推進事業 介護サービス実施記録の電子的共有を可能とする情報システム基盤に関する調査研究報告書」(介護サービス実施記録システム普及コンソーシアム 2005年3月)
- 「標準的な福祉介護用語群の開発・活用事業報告書」(財団法人医療情報システム開発センター 2005年3月)

Basic Analysis of Terms and Records in Welfare and Care Practice — Methods and technique of analysis for records of service provision —

IKUTA Masayuki *

Abstract: Records of welfare and care practice have been made with much cost and labor, however, most of them are used only to a limited extent and rarely utilized systematically and routinely. Recently, the situation surrounding welfare and care has been changing drastically and tightening, and now, a movement to make positive use of records is enhancing to improve QOL of service users and to promote effective and efficient service provision. To investigate the possibility of record utilization, we collected a large amount of records of service provision in care practice, analyzed records morphologically using computer and interpreted actual records based on term usage as well as evaluated analysis methods and technique, in addition, discussed problems for future study.

Keywords: welfare, care, welfare information, record of service provision, computer, morphological analysis, word and term standardization, varying notation, synonym

* Professor, Faculty of Social Sciences, Ritsumeikan University